# Technical Report: Computationally Sound Secrecy Proofs by Mechanized Flow Analysis*

Michael Backes[1], Peeter Laud[2]

[1] Saarland University
[2] Tartu University

January 21, 2010

**Abstract.** We present a novel approach for proving secrecy properties of security protocols by mechanized flow analysis. In contrast to existing tools for proving secrecy by abstract interpretation, our tool enjoys cryptographic soundness in the strong sense of blackbox reactive simulatability/UC which entails that secrecy properties proven by our tool are automatically guaranteed to hold for secure cryptographic implementations of the analyzed protocol, with respect to the more fine-grained cryptographic secrecy definitions and adversary models.
Our tool is capable of reasoning about a comprehensive language for expressing protocols, in particular handling symmetric encryption and asymmetric encryption, and it produces proofs for an unbounded number of sessions in the presence of an active adversary. We have implemented the tool and applied it to a number of common protocols from the literature.

## 1 Introduction

Security proofs of cryptographic protocols are known to be difficult and the automation of such proofs has been studied soon after the first protocols were developed. From the start, the actual cryptographic operations in such proofs were idealized into so-called Dolev-Yao models, following [2–4], e.g., see [5–10]. This idealization simplifies proof construction by freeing proofs from cryptographic details such as computational restrictions, probabilistic behavior, and error probabilities. Conducting secrecy proofs by typing based on these abstractions has shown to be a particularly salient technique as it allowed for elegant and fully automated proofs, often even for an unbounded number of sessions.

A type system was recently presented in [11] that combines the conciseness of language-based reasoning in Dolev-Yao models with strong computational soundness guarantees, i.e., if an abstract protocol typechecks then its cryptographic realization provably keeps the quantities handed to it by the protocol users (payload data) secret in the computational sense. Such computational soundness guarantees of abstract proofs have recently been identified as central for gaining trustworthy guarantees of

* An earlier version of this work appeared in [1].

security protocols: the computational model strives for stronger, more fine-grained security notions and furthermore considers a more realistic adversary that is allowed to perform arbitrary bitstring manipulations as long as they can be performed in probabilistic polynomial-time. However, despite being the first type system that allows for abstract, computationally sound reasoning under active attacks, the major drawback of [11] was that type inference was not considered. As a consequence, this work did not entail an automated procedure for analyzing secrecy aspects of cryptographic protocols with cryptographic soundness guarantees, which arguably is the central goal of unifying the advantages of both approaches.

We remedy this shortcoming by presenting a mechanized approach for soundly proving secrecy of payload data in cryptographic protocols by analysing the possible flows of data during the execution of the protocol. Our approach is capable of reasoning about a comprehensive language for expressing protocols, in particular handling symmetric encryption and asymmetric encryption, allows for more precise analyses compared with the type system of [11], is fully automated, and produces proofs for an unbounded number of sessions in the presence of an active adversary.

Our results (and the one of [11] as well) rely on a variant of the Dolev-Yao model of Backes, Pfitzmann, and Waidner, henceforth called the BPW model, which has been shown to be computationally sound in the strong sense of of blackbox reactive simulatability (BRSIM). The security notion of BRSIM means that one system (here, the cryptographic realization) can be plugged into arbitrary protocols instead of another system (here, the BPW model) [12–15]. While first security proofs of several common protocols have been hand-crafted using the BPW model [16–21], recent work has shown that the BPW model is accessible to theorem proving techniques as well [22]. Our work shows that soundly proving secrecy properties in a fully automated manner is possible using the BPW model, and it identifies cryptographically sound secrecy by typing as a promising direction for future work in general. In particular, our line complements the large number of existing works that aims at establishing computational soundness of Dolev-Yao models without considering secrecy by typing, cf. the section on related work for more details.

The analysis presented in this paper builds on the spi-calculus-style language, its deterministic semantics and the corresponding type system from [11] and is inspired by methods from control flow analysis. It works by collecting for each defined variable at each protocol point the possible shapes of terms that this variable may point to, including the possible creation points of the atomic subterms. The same information is also collected for channels between participants for encryption keys, thus yielding information which terms may be communicated over which channel, and which terms may be encrypted with which keys by honest participants, respectively. Finally, the same abstraction is also collected for terms that the adversary may learn during the run of the protocol.

There are a couple of noteworthy points. First, all inputs from the adversary are modeled using a single abstract value, thus freeing the analyser from the necessity to model every new term that the adversary may construct. Instead, we consider explicit rules for decomposing this abstract value, i.e., the adversary's input, which allows us to keep the description of the adversary's knowledge finite. Secondly, parts of the protocol

statically following a public-key decryption are analysed twice — once assuming that the ciphertext was created by an honest participant and once assuming that it was created by the adversary. The distinction of these two cases (which was already present in [23] and also in [11]) is important for the precision of the analyser. Thirdly, we collect not only the possible values of variables but also relationships between them. Whenever certain operations restrict the set of possible values of some variables, we exploit these recorded relationships in order to restrict the set of values of related variables as well. This collection of relationships is reminiscent of shape analysis [24], although our task is considerably simpler here than a full shape analysis because we do not have destructive updates. We record the relationships between variables by collecting a set of constraints that their abstractions must satisfy.

Our prover (consisting of constraint generator and solver) has been implemented in O'Caml and can be downloaded at `http://www.ut.ee/~peeter_l/research/brsiman`.

### 1.1 Related Work

Early work on linking Dolev-Yao-style symbolic models and cryptography [25–28] only considered passive attacks, and therefore cannot make general statements about protocols.

The security notion of BRSIM was first defined generally in [12], based on simulatability definitions for secure (one-step) function evaluation. It was extended in [14, 13], the latter with somewhat different details and called UC (universal composability), and has been widely applied to prove individual cryptographic systems secure and to derive general theoretical results. In particular, BRSIM/UC allows for plugging one system into arbitrary protocols instead of another system while retaining essentially arbitrary security properties [12, 13, 39].

A cryptographic justification of a Dolev-Yao model in the sense of BRSIM/UC was first given in [30]. Extensions of this BRSIM/UC result to more cryptographic primitives were presented in [31–33] and used in protocol proofs [16–21]. General theorems on property preservation through the BRSIM notion imply that the same Dolev-Yao model and realization also fulfill some other soundness notions [12, 34–36, 34, 37, 38], and further soundness results specific to this Dolev-Yao model and realization were proved in [39]. Some later papers [40–42] considered to what extent restrictions to weaker security properties or less general protocol classes allow simplifications compared with [30]: Laud [41] has presented cryptographic foundations for a Dolev-Yao model of symmetric encryption but specific to certain confidentiality properties where the surrounding protocols are restricted to straight-line programs. Warinschi et al. [40, 43] have presented cryptographic underpinnings for a Dolev-Yao model of public-key encryption, yet for a restricted class of protocols and protocol properties that can be analyzed using this primitive. Baudet, Cortier, and Kremer [44] have established the soundness of specific classes of equational theories in a Dolev-Yao model under passive attacks. Canetti and Herzog [42] have shown that a Dolev-Yao-style symbolic analysis can be conducted using the framework of universal composability for a restricted class of protocols, namely mutual authentication and key exchange protocols with the

additional constraint that the protocols must be expressible as loop-free programs using public-key encryption as their only cryptographic operation. We stress that none of these works build on type inference for proving secrecy properties of security protocols. Since computational soundness has become a highly active line of research, we exemplarily list further recent results in this area without going into further details [45–50, 33, 51].

The work that comes closest to our work is the work of Laud [11] who designed a type system for proving secrecy aspects of security protocols based on the BPW model. He shows that if an abstract protocol typechecks in his system, then its cryptographic realization provably keeps the quantities handed to it by the protocol users secret in the computational sense. The proof of this fact exploits the BRSIM/UC soundness result of [30, 32, 29] for carrying over symbolic proofs of secrecy in the BPW model to the actual cryptographic realization, similar to the present paper. However, type inference has not been implemented yet in this paper so that the paper did not entail a mechanized procedure for soundly proving secrecy aspects of security protocols.

Efforts are also under way to formulate syntactic calculi with a probabilistic, polynomial-time semantics, including approaches based on process algebra [52, 53], security logics [54, 55] and cryptographic games [56]. In particular, Datta et al. [55] have proposed a promising logical deduction system to prove computational security properties. We are not aware of any implementations of these frameworks, except for Blanchet's [56], who has recently presented an automated tool for proving secrecy properties of security protocols based on transforming cryptographic games. This line of work is orthogonal to the work of justifying Dolev-Yao models, which offer a higher level of abstractions and thus much simpler proofs where applicable, so that proofs of larger systems can be automated.

Let us also mention some of the work in the area of type systems for cryptographic protocol analysis. The first type system of this kind was proposed by Abadi [57], which could be used for verifying the secrecy of payloads or nonces in the protocols using only symmetric encryption. This type system, as well as all the remaining ones that we describe work in the Dolev-Yao model. The type system was extended to cope with asymmetric encryption by Abadi and Blanchet [23]. Abadi and Blanchet [58] further generalized this type system to handle *generic* cryptographic primitives. The type system of Abadi has also been extended by Gordon and Jeffrey [59–61] to check for integrity properties. Finally, a static program analysis [62] and a type system [63] exist that work directly in the computational model, handling programs containing symmetric encryption, both for passive adversaries only.

Abstract interpretation, which is in most cases automatable using data flow analysis, has also been considered for the analysis of cryptographic protocols within the Dolev-Yao model. See for example [64] and the references contained therein.

### 1.2 Structure of the paper

We start by describing our (machine-based) execution model and the language used to program these machines for expressing security protocols in Sec. 2. We continue in Sec. 3 with the description of the analysis. In particular, we give the correctness theorem stating under which conditions the results of the abstract analysis entail computational

security of a cryptographic protocol. Sec. 4 describes the implementation of our tool and its applicability to common security protocols from the literature. In Sec. 5 we give the main technical lemma, similar to subject reduction, used to prove the previously given correctness theorem.

## 2 Execution Model

We use the same setup of a system as in [11]. In short, the BPW model (sometimes also called abstract cryptographic library in the following corresponding to the original title of [30]) for $n$ honest users is implemented by a machine $\mathcal{TH}_n$ which has input ports $\mathsf{in}_{\mathsf{u}_i}?$ to receive commands from the $i$-th user, output ports $\mathsf{out}_{\mathsf{u}_i}!$ to return the results of commands and (handles of) received messages, ports $\mathsf{in}_{\mathsf{a}}?$ and $\mathsf{out}_{\mathsf{a}}!$ for the communication with the adversary, and a database of terms. The database records the structure of messages and the knowledge of messages by the parties ($n$ users and the adversary). The users and the adversary access messages through *handles*, the transmission of messages involves the translation of handles. The possible commands are the construction, taking apart, and sending of messages. The protocol logic for the $i$-th user is implemented by a machine $\mathsf{P}_i$ that connects to the ports $\mathsf{in}_{\mathsf{u}_i}?$ and $\mathsf{out}_{\mathsf{u}_i}!$ and offers the ports $\mathsf{pin}_{\mathsf{u}_i}?$ and $\mathsf{pout}_{\mathsf{u}_i}!$ to the user through which it may send and receive data. An execution step of a machine $\mathsf{P}_i$ consists of receiving a message (either from $\mathcal{TH}_n$ or the user), performing some computations on the terms, and optionally sending a message. The machines $\mathsf{P}_i$ are programmed in a language resembling the spi-calculus, defined below.

$$
\begin{array}{rl}
e ::= & n \mid \underline{\mathsf{keypair}}^{\ell} \mid \mathsf{store}(x) \\
\mid & x \mid \mathsf{retrieve}(x) \mid \mathsf{list}(x_1, \ldots, x_k) \\
& \mid \underline{\mathsf{pubkey}}(x) \mid \underline{\mathsf{pubenc}}^{\ell}(x_{\mathrm{k}}, x_{\mathrm{t}}) \\
& \mid \underline{\mathsf{privenc}}^{\ell}(x_{\mathrm{k}}, x_{\mathrm{t}}) \mid \underline{\pi_i^j}(x) \\
& \mid \mathsf{gen\_symenc\_key}(i)^{\ell} \mid \underline{\mathsf{pubdec}}(x_{\mathrm{k}}, x_{\mathrm{t}}) \\
& \mid \underline{\mathsf{privdec}}(x_{\mathrm{k}}, x_{\mathrm{t}}) \mid \mathsf{gen\_nonce}^{\ell}
\end{array}
$$

$$
\begin{array}{rl}
SIP ::= & \mathbf{receive}_c^{\ell}[x_{\mathrm{p}}](x) \\
IP ::= & SIP \mid !SIP \\
I ::= & IP.P \\
I^* ::= & \mathbf{0} \mid I \mid I^*
\end{array}
\qquad
\begin{array}{rl}
P ::= & I^* \mid \mathcal{II} \\
\mid & \mathbf{send}_c[x_{\mathrm{p}}](x).I^* \\
\mid & \mathbf{let}^{\ell}\ x := e\ \mathbf{in}\ P \\
& \qquad \mathbf{else}\ P' \\
\mid & \mathbf{if}^{\ell}\ x = x'\ \mathbf{then}\ P \\
& \qquad \mathbf{else}\ P'
\end{array}
$$

Here $x$-s are variables, $e$-s are expressions, $I$-s are input processes, $P$-s are output processes, and $\ell$-s are labels for program points and expressions of interest. No label may occur twice in the protocol text, nor can a variable be defined twice or used before being defined. The language contains public-key and symmetric-key encryption as the cryptographic primitives (as well as nonces). A public and secret key pair is created by the expression $\underline{\mathsf{keypair}}$, the public key is extracted by $\underline{\mathsf{pubkey}}$. A *level* $i$ is associated with each symmetric key to prevent encryption cycles (and make the proof relating $\mathcal{TH}_n$ and its concrete implementation go through); a symmetric key may only encrypt keys of lower level. The store- and retrieve-expressions are used to convert payloads (data that

can be communicated with the user) to handles and back. The expression $\pi_i^j(x)$ extracts the $i$-th component from the list of length $j$ pointed to by $x$. In $\mathbf{receive}_c[x_\mathrm{p}](x)$ and $\mathbf{send}_c[x_\mathrm{p}](x)$, the variable $x$ is the message and $x_\mathrm{p}$ is the identity of the other party. The channel for the message is given by the constant *abstract channel* $c$. An abstract channel is used to group messages sent between protocol participants, as well as between the protocol user and participant (although the abstract channel does not alone determine the sender and the receiver of a message). Furthermore, the set of abstract channels $\mathbf{Chan}$ is partitioned into four parts, denoted $\mathbf{Chan}_\mathrm{x}$, where $\mathrm{x} \in \{\mathsf{s}, \mathsf{a}, \mathsf{i}, \mathsf{u}\}$. If a message is sent on an abstract channel from $\mathbf{Chan}_\mathsf{s}$ [resp. $\mathbf{Chan}_\mathsf{a}$, $\mathbf{Chan}_\mathsf{i}$] then it means that the message travels between protocol participants over a secure (resp. authentic, insecure) channel. If a message is sent on an abstract channel from $\mathbf{Chan}_\mathsf{u}$ then it travels between the protocol user and the protocol participant (i.e. over one of the concrete channels $\mathsf{pin}_{\mathsf{u}_i}$ or $\mathsf{pout}_{\mathsf{u}_i}$). The variables $x$ and $x_\mathrm{p}$ are bound in a $\mathbf{receive}$-statement. The variable $x$ is also bound in the default-branch of a $\mathbf{let}$-statement, but not in the else-branch, which is taken upon a failure of evaluating $e$.

The internal state of an inactive (i.e. not currently running) $\mathsf{P}_i$ consists of a list of input processes together with their execution environments, giving values to already defined variables. The "program" (or initial state) of each $\mathsf{P}_i$ is a list of input processes. An active $\mathsf{P}_i$ additionally contains the received message (together with the apparent sender and the name of the channel it was received on) and the currently running (output) process (together with its environment). When $\mathsf{P}_i$ receives the message, it is handed over to the first input process $(!)\mathbf{receive}_c[x_\mathrm{p}](x).P$ with matching channel name $c$ in its list of processes. The variables $x$ and $x_\mathrm{p}$ are bound to the message and the apparent sender and the process executes until it has become $\mathcal{II}$ or a list of input processes $I^*$. The value $\mathcal{II}$ means rejecting the message — the list of input processes of $\mathsf{P}_i$ is not changed, the currently executing process and its environment are discarded (thereby forgetting all references to any new terms that may have been created since receiving that message) and the message is handed over to the next input process with the matching channel name in the list of input processes of $\mathsf{P}_i$. When a process accepts the message, it executes until it has become a list of input processes $I^*$. All processes in this list $I^*$, together with the environment of the output process, are put to the list of input processes of $\mathsf{P}_i$ instead of or in addition of (depending on the presence of replication) the original process. When no process accepts the message, it is simply lost.

### Security

The security property we are interested in is the *secrecy of payloads* [29]. We also considered the same property in [11] and our treatment here does not differ from that. In short, we want the system implementing the protocol (consisting of the machines $\mathsf{P}_1, \ldots, \mathsf{P}_n$ and $\mathcal{TH}_n$) to retain the secrecy of any payloads *handed to it by the users* over the ports $\mathsf{pin}_{\mathsf{u}_i}?$. The secrecy of payloads means that the user and the adversary together cannot figure out whether the system implementing the protocol is really computing with the values received from the user or with some other values. In the definition of payload secrecy, there is a scrambler / descrambler inserted between the user and the system implementing the protocol. If it is turned on, it replaces all messages from the user to the system (i.e. all payloads) with random values; and replaces these random

values, if they are sent from the system to the user, with the values received from the user again. The user (together with the adversary) has to guess whether the scrambler / descrambler is turned on. If the user cannot guess then the payloads are kept secure in the system — they do not flow from the user through the system to the adversary. A precise definition can be found in [29] and a concise description in [11]. In [11] the following five properties were stated to be sufficient for the secrecy of payloads and for the simulatability of the machine $\mathcal{TH}_n$:

(I) the bit-strings that the machines $\mathsf{P}_i$ receive from the ports $\mathsf{pin}_{\mathsf{u}_i}?$ do not affect the control flow of $\mathsf{P}_i$, i.e. this data is not used in the **if**-statements;
(II) the machines $\mathsf{P}_i$ may pass the bit-strings received from the user to the cryptographic library only in store-commands;
(III) the terms resulting from these store-commands will not become available to the adversary, i.e. the adversary does not get handles for these terms.
(IV) symmetric keys of order $i$ only encrypt terms of order less than $i$ (note that symmetric keys created by the adversary have no order and are thereby not restricted by this condition);
(V) if a symmetric key unknown to the adversary (i.e. the adversary does not have a handle to it) is used for encryption then this key will never become known to the adversary.

The analysis presented in this paper verifies that these five properties hold.

A different secrecy property was also considered in [29] — the secrecy of keys generated by the system during the protocol run. We do not consider this property here, although a corresponding list of sufficient properties would not be difficult to fix (key secrecy can be considered a simpler property than payload secrecy); and these properties could also be verified by our analysis. However, this remains future work.

## 3 Analysis

We set up a constraint system whose solutions upper-approximate the values flowing through the protocol. A constraint system consists of two main parts — constraint variables and the set of constraints. A constraint variable is just an identifier together with an associated domain (an upper semilattice) of possible values. A constraint is a statement of the form $E \leq C$ where $C$ is a constraint variable and $E$ is a monotone expression over constraint variables. A solution of a constraint system is a valuation, mapping each constraint variable to a value in its domain, such that all constraints are satisfied.

We will prove that any solution to our constraint system will be a safe approximation of any possible protocol run. To get the best precision, we are interested in the least solution. There are well-known methods for finding the least solutions for constraint systems with monotone constraints.

### 3.1 Abstract Domain

The possible values of protocol variables are abstracted by sets of the following abstract values $AV$. The sets of abstract values are used as domains for certain constraint

variables below.

$$AV ::= AV_I \mid AV_H \mid \mathsf{seckey}(\ell, b) \qquad AV_I = \mathsf{X_P} \mid \mathsf{X_S}$$

$$\begin{array}{rl} AV_H ::= & \mathsf{pubenc}(AV_H, AV_H, \ell, b) \mid \mathsf{nonce}(\ell, b) \\ \mid & \mathsf{symenc}(AV_H, AV_H, \ell, b) \mid \mathsf{AnyPubVal} \\ \mid & \mathsf{symkeyname}(\ell, b) \mid \mathsf{pubkey}(\ell, b) \\ \mid & (AV_H, \ldots, AV_H) \mid \mathsf{store}(AV_I) \\ \mid & \mathsf{symkey}(i, \ell, b) \end{array} \tag{1}$$

Here $AV_I$ contains the possible abstractions of payloads — they may be either public ($\mathsf{X_P}$) or secret ($\mathsf{X_S}$). The addresses of the communication partners (variable $x_\mathrm{p}$ in **send**- and **receive**-commands) are public. Data received from the protocol users are secret. The terms $AV_H$ are the possible abstractions of terms in the database of $\mathcal{TH}_n$. They should be mostly self-descriptive. The arguments $\ell$ refer to program points (labels at expressions) where these values have been created. The arguments $b$ also resemble program points — we have mentioned before that we analyse the parts of the protocol following a public-key decryption twice — once assuming that the ciphertext was generated by a protocol participant and once assuming that it was generated by the adversary. Hence, if $n$ public-key decryptions occur before the program point $\ell$ then this point really counts as $2^n$ different program points for the analysis. If $\ell$ is a program point following $n$ decryptions then $b$ is a bit-string of length $n$ where $i$-th bit records the assumed creator of the $i$-th decrypted ciphertext (1 — some honest participant; 0 — the adversary). We call $b$ the *decryption context*.

The argument $i$ in $\mathsf{symkey}(i, \ell, b)$ records the level of the symmetric key. The abstract value $\mathsf{symkeyname}(\ell, b)$ corresponds to the identities of the symmetric keys created at the program point $\ell$ (with the decryption context $b$). According to $\mathcal{TH}_n$, the adversary is able to find the identities of symmetric keys from the ciphertexts created with them. The abstract value $\mathsf{AnyPubVal}$ denotes any value that the adversary knows and may have constructed. All other $AV_H$ denote values constructed by protocol participants. The secret decryption keys $\mathsf{seckey}(\ell, b)$ are not listed as a possible case for $AV_H$ because $\mathcal{TH}_n$ puts severe restrictions on their use — they may only be used for decrypting ciphertexts; they cannot appear as subterms of more complex terms.

### 3.2 Constraint Variables

Given a protocol $\wp$ with its set of labels, we introduce the following constraint variables.

*First,* $\mathbf{S}^b_\ell$ for all statement labels $\ell$ occurring in the protocol (here we only consider the labels of **if**-, **let**- and **receive**-statements, not the labels occurring in expressions). Here $b$ is a bit-string whose length equals the number asymmetric decryption operations that occur in the protocol before and including the point labeled with $\ell$. Hence for a program point $\ell$ that is preceded by $n$ asymmetric decryptions we have $2^n$ different variables $\mathbf{S}^b_\ell$.

Let $\mathbf{Var}^\circ_\ell$ be the set of variables defined before the protocol point $\ell$. Let $\mathbf{Var}^\bullet_\ell$ be the union of $\mathbf{Var}^\circ_\ell$ with the set of protocol variables that are assigned a value at $\ell$ (depending on whether $\ell$ labels an **if**-, **let**- or **receive**-statement, this set has 0, 1 or 2 elements). The possible values for $\mathbf{S}^b_\ell$ are mappings from $\mathbf{Var}^\bullet_\ell$ to sets of abstract

values $AV$. These mappings are ordered pointwise, with the sets of abstract values $AV$ ordered by subset inclusion.

The variable $\mathbf{S}_\ell^b$ records the possible values of protocol variables after a successful completion of the operation at program point $\ell$. A $\mathbf{let}$- or $\mathbf{if}$-statement is successful if the default-/true-branch was taken. A $\mathbf{receive}$-statement is always successful.

*Second,* $\mathbf{R}_\ell^b$ for all statement labels $\ell$ occurring in the protocol and $b$ having the same possible values as for $\mathbf{S}_\ell^b$ (and ordered the same way). These constraint variables are introduced to ease the presentation of the constraint system. Namely, the handling of a statement (if it succeeds) proceeds in two steps: first the constraints giving the abstraction(s) of the newly defined variable(s) are evaluated, followed by the evaluation of constraints describing the relationships between the values of different variables. The constraint variable $\mathbf{S}_\ell^b$ contains the result of these two steps. The constraint variable $\mathbf{R}_\ell^b$ contains the result of the first step only.

*Third,* $\mathbf{C}_c$ for all abstract channels $c \in \mathbf{Chan_s} \cup \mathbf{Chan_a}$ occurring in the protocol. The possible values of these variables are sets of abstract values $AV_H$, ordered by subset inclusion. These variables will record an abstraction of the possible messages sent over the abstract channel $c$.

*Fourth,* $\mathbf{P}$. This will record the values that the adversary knows. The possible values of this variable are sets of abstract values $AV$, ordered by subset inclusion.

*Fifth,* $\mathbf{E}_\ell^b$ for a label $\ell$ occurring at a key generation. This set records all abstract values that are encrypted with the key generated at $\ell$ for the preceding asymmetric decryption results described by $b$. The bit-string $b$ has the same meaning as for the variables $\mathbf{S}_\ell^b$ (the point of interest is the occurrence of $\ell$ in the protocol). The possible values of these variables are sets of abstract values $AV_H$, ordered by subset inclusion.

*Sixth,* $\mathbf{L}_{\ell,\mathsf{true}}^b$ and $\mathbf{L}_{\ell,\mathsf{false}}^b$ for labels $\ell$ at *let*- and *while*-statements. They denote whether the true- (default-) and false-branch of the statement are alive or not. The bit-string has the same meaning as before. The possible values of these variables are $\mathsf{false}$ and $\mathsf{true}$, ordered by $\mathsf{false} \leq \mathsf{true}$.

### 3.3 Constraints

There are two sources for constraints — the protocol and the adversary. The first describes the movement of values during the computations performed by the protocol, while the second describes the capabilities of the adversary in decomposing messages. This second set of constraints, given in Fig. 1 is quite straightforward: The first two constraints are obvious — the adversary can retrieve stored payloads and decompose lists. The third constraint states that the adversary can decrypt a symmetric encryption if it has the key. The relation $\cong_{\mathbf{P}}$ relates two abstract values if the sets of terms they correspond to may intersect. Because the meaning of $\mathsf{AnyPubVal}$ depends on the adversary's knowledge, this relation must also depend on it. The relation $\cong_{\mathbf{P}}$ is the least reflexive, symmetric and structure-respecting relation on abstract values that satisfies the conditions given in Fig. 2. The fourth constraint for the adversary's capabilities states that if the public key used for public encryption may have been created by the adversary (which means that the secret key was also created by the adversary) then the adversary may find out the plaintext. The fifth and sixth constraints state that the adversary is capable of determining the identity of the key used to produce the ciphertext.

$$\mathsf{store}(AV) \in \mathbf{P} \Rightarrow AV \in \mathbf{P}$$
$$(AV_1, \ldots, AV_j) \in \mathbf{P} \Rightarrow AV_i \in \mathbf{P}$$

$$\mathsf{symenc}(AV_k, AV_t, \ell, b) \in \mathbf{P} \Rightarrow (\exists AV' \in \mathbf{P} : AV_k \cong_{\mathbf{P}} AV') \Rightarrow AV_t \in \mathbf{P}$$

$$\mathsf{pubenc}(\mathsf{AnyPubVal}, AV_t, \ell, b) \in \mathbf{P} \Rightarrow AV_t \in \mathbf{P}$$
$$\mathsf{pubenc}(AV_k, AV_t, \ell, b) \in \mathbf{P} \Rightarrow AV_k \in \mathbf{P}$$

$$\mathsf{symenc}(\mathsf{symkey}(i, \ell, b), AV_t, \ell', b') \in \mathbf{P} \Rightarrow \mathsf{symkeyname}(\ell, b) \in \mathbf{P}$$

$$\{\mathsf{X_P}, \mathsf{AnyPubVal}\} \subseteq \mathbf{P}$$

**Fig. 1.** Constraints describing the adversary's power

$$AV \in \mathbf{P} \Rightarrow AV \cong_{\mathbf{P}} \mathsf{AnyPubVal}$$
$$\mathsf{store}(\mathsf{X_P}) \cong_{\mathbf{P}} \mathsf{AnyPubVal}$$

$$\left(\forall i : AV_i \cong_{\mathbf{P}} AV'_i\right) \wedge (AV'_1, \ldots, AV'_j) \cong_{\mathbf{P}} \mathsf{AnyPubVal} \Rightarrow (AV_1, \ldots, AV_j) \cong_{\mathbf{P}} \mathsf{AnyPubVal}$$

$$(\mathsf{AnyPubVal}, \ldots, \mathsf{AnyPubVal}) \cong_{\mathbf{P}} \mathsf{AnyPubVal}$$

$$AV_k \cong_{\mathbf{P}} AV'_k \wedge AV_t \cong_{\mathbf{P}} AV'_t \wedge \mathsf{pubenc}(AV'_k, AV'_t, \ell, b) \in \mathbf{P} \Rightarrow \mathsf{pubenc}(AV_k, AV_t, \ell, b) \cong_{\mathbf{P}} \mathsf{AnyPubVal}$$

$$AV_k \cong_{\mathbf{P}} AV'_k \wedge AV_t \cong_{\mathbf{P}} AV'_t \wedge \mathsf{symenc}(AV'_k, AV'_t, \ell, b) \in \mathbf{P} \Rightarrow \mathsf{symenc}(AV_k, AV_t, \ell, b) \cong_{\mathbf{P}} \mathsf{AnyPubVal}$$

$$\mathsf{X_S} \cong_{\mathbf{P}} \mathsf{X_P}$$

**Fig. 2.** The "possibly equal" relation $\cong_{\mathbf{P}}$

For asymmetric encryption, this identity is the public key itself, while for symmetric encryption, it is the symkeyname. Finally, the public values $\mathsf{X_P}$ and AnyPubVal may be known to the adversary.

The set of constraints generated by an input or output process $P$ is given by the mapping $\langle\!\langle\!\langle P \rangle\!\rangle\!\rangle$ that we are going to define below. For defining it, we also define the following mappings.

- $\langle\!\langle e \rangle\!\rangle(\mathbf{I}, b)$ gives the set of abstract values for the result of evaluating the expression $e$ when the decryption context (after evaluating $e$) is $b$ and the abstractions of already defined variables are given by the mapping $\mathbf{I}$.
- $\langle\!\langle e \rangle\!\rangle_{\mathrm{s}}(\mathbf{I})$ and $\langle\!\langle e \rangle\!\rangle_{\mathrm{f}}(\mathbf{I})$ give some necessary conditions for the evaluation of $e$ to succeed or fail.
- $\langle\!\langle e \rangle\!\rangle_{\mathcal{E}}(\mathbf{I}, \mathbf{L})$ gives the set of constraints for the variables $\mathbf{E}^b_\ell$, as generated by $e$. Here $\mathbf{L}$ is a boolean showing whether this expression is live code.
- $[\![x := e]\!](b, \mathbf{X}, y)$, where $\mathbf{X}$ gives the abstractions of variables defined before the assignment $x := e$, and $y$ is either $x$ or a variable occurring in $e$ gives a set of

abstract values that certainly abstracts the value of $y$ after the successful evaluation of $e$. The mapping $\lfloor\!\lfloor x := e \rfloor\!\rfloor$ is used to collect the relationships between values of variables.

The relationships between variables allow us to make the analysis more precise — they bound the abstractions of values of variables from above. When combining several of these upper bounds, we have to form their greatest lower bound. While the least upper bound of two sets of abstract values may be just the union of sets, the greatest lower bound cannot be simply the intersection because when two sets of abstract values $\mathbf{A}$ and $\mathbf{B}$ are both valid abstractions of some concrete value then we want their greatest lower bound $\mathbf{A} \mathbin{\dot\cap} \mathbf{B}$ be a valid abstraction of that value as well. But certain concrete values may correspond to several different abstract values, for example a nonce that has become known to the adversary may occur in our abstractions either as $\mathsf{nonce}(\ell, b)$ for some $\ell$ and $b$ or as $\mathsf{AnyPubVal}$.

For defining $\mathbin{\dot\cap}$, we first define a *partial* binary operation $\sqcap$ on abstract values as the smallest (i.e. defined for as few arguments as possible) idempotent symmetric structure-preserving operation that satisfies $AV_H \sqcap \mathsf{AnyPubVal} = AV_H$ for any abstract value $AV_H$ defined in (1). Now we can just define $\mathbf{A} \mathbin{\dot\cap} \mathbf{B} = \{AV \sqcap AV' \mid AV \in \mathbf{A}, AV' \in \mathbf{B}\}$. We also define $\mathbf{A} \mathbin{\dot\subseteq} \mathbf{B}$ iff $\mathbf{A} \mathbin{\dot\cap} \mathbf{B} = \mathbf{A}$.

The mappings $\langle\!\langle e \rangle\!\rangle$, $\langle\!\langle e \rangle\!\rangle_{\mathrm{s}}$, $\langle\!\langle e \rangle\!\rangle_{\mathrm{f}}$ and $\langle\!\langle e \rangle\!\rangle_{\mathcal{E}}$ are given in Figures 3 and 4. If we have left out the definition of $\langle\!\langle e \rangle\!\rangle_{\mathrm{s}}$ or $\langle\!\langle e \rangle\!\rangle_{\mathrm{f}}$ for some $e$ then it is $\mathsf{true}$. If we have left out the definition of $\langle\!\langle e \rangle\!\rangle_{\mathcal{E}}$ for some $e$ then it is $\emptyset$. In Fig. 4, $\mathsf{L}_a$ [resp. $\mathsf{L}_s$] denotes the set of all labels $\ell$ occurring in the protocol in the positions $\underline{\mathsf{keypair}}^\ell$ [resp. $\mathsf{gen\_symenc\_key}(i)^\ell$]. In Fig. 3 we can see the special treatment of $\mathsf{AnyPubVal}$ — for example, payloads can be extracted from it and projections can be taken. The result is still a public value. During encryption, $\mathsf{AnyPubVal}$ may serve as the encryption key (of course, such ciphertexts can be decrypted by the adversary). During decryption, when the ciphertext is $\mathsf{AnyPubVal}$, we use the variables $\mathbf{E}^b_\ell$ to determine the possible plaintexts.

The distinction between participant-generated and adversarially generated ciphertexts in public-key decryption can be seen in two definitions for $\langle\!\langle \underline{\mathsf{pubdec}}(x_k, x_t) \rangle\!\rangle$. First of them assumes that the ciphertext is generated by some protocol participant, while the second assumes that the adversary is the source of the ciphertext. Both of these cases are also present in symmetric decryption, but they have been joined together, so that the analysis does not handle them separately.

The relationships between newly defined and existing variables are given by $\lfloor\!\lfloor x := e \rfloor\!\rfloor(b, \mathbf{X}, y)$, defined in Fig. 5. Recall that it gives for a variable $y$ a set of abstract values that is guaranteed to abstract its concrete value. If the definition is missing for some $\lfloor\!\lfloor x := e \rfloor\!\rfloor(b, \mathbf{X}, y)$ in Fig. 5, it is equal to $\mathbf{X}(y)$ (i.e. no precision is gained). In Fig. 5, the message constructor $\mathsf{C}$ may be one of $\mathsf{list}$, $\underline{\mathsf{pubenc}}^\ell$ or $\underline{\mathsf{privenc}}^\ell$. The corresponding abstract value constructor $C$ is then either the list constructor, $\mathsf{pubenc}(\cdot, \cdot, \ell, b)$ or $\mathsf{privenc}(\cdot, \cdot, \ell, b)$.

The usage of $\lfloor\!\lfloor x := e \rfloor\!\rfloor$ may become clearer when we look at the constraints generated by processes. This generation is done by $\langle\!\langle\!\langle P \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C})$ where $\mathbf{I}$ is a constraint variable describing the protocol state before the execution of the process $P$, the bitstring $b$ is the current decryption context, the variable $\mathbf{L}$ denotes whether this process is alive, and $\mathcal{C}$ is a set of constraints of the form $\mathbf{X}(x) \mathbin{\dot\subseteq} E$ where $E$ is a monotone

$$\langle\!\langle n \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{X_P}\} \qquad \langle\!\langle x \rangle\!\rangle(\mathbf{I}, b) = \mathbf{I}(x)$$
$$\langle\!\langle \underline{\mathsf{keypair}}^{\ell} \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{seckey}(\ell, b)\}$$
$$\langle\!\langle \mathsf{store}(x) \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{store}(AV) \mid AV \in \mathbf{I}(x)\}$$

$$\langle\!\langle \mathsf{retrieve}(x) \rangle\!\rangle(\mathbf{I}, b) = \{AV \mid \mathsf{store}(AV) \in \mathbf{I}(x)\} \cup \{\mathsf{X_P} \mid \mathsf{AnyPubVal} \in \mathbf{I}(x)\}$$

$$\langle\!\langle \mathsf{list}(x_1, \ldots, x_k) \rangle\!\rangle(\mathbf{I}, b) = \{(AV_1, \ldots, AV_k) \mid AV_i \in \mathbf{I}(x_i)\}$$
$$\langle\!\langle \mathsf{gen\_symenc\_key}(i)^{\ell} \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{symkey}(i, \ell, b)\}$$

$$\langle\!\langle \underline{\pi_i^j}(x) \rangle\!\rangle(\mathbf{I}, b) = \{AV_i \mid (AV_1, \ldots, AV_k) \in \mathbf{I}(x)\} \cup \{\mathsf{AnyPubVal} \mid \mathsf{AnyPubVal} \in \mathbf{I}(x)\}$$

$$\langle\!\langle \underline{\mathsf{pubkey}}(x) \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{pubkey}(\ell, b) \mid \mathsf{seckey}(\ell, b) \in \mathbf{I}(x)\}$$
$$\langle\!\langle \mathsf{gen\_nonce} \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{nonce}(\ell, b)\}$$

$$\langle\!\langle \underline{\mathsf{pubenc}}^{\ell}(x_k, x_t) \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{pubenc}(AV_k, AV_t, \ell, b) \mid AV_k \in \mathbf{I}(x_k), AV_t \in \mathbf{I}(x_t), AV_k = \mathsf{pubkey}(\ldots)\} \cup \{\mathsf{pubenc}(\mathsf{AnyPubVal}, AV_t, \ell, b) \mid \mathsf{AnyPubVal} \in \mathbf{I}(x_k), AV_t \in \mathbf{I}(x_t)\}$$

$$\langle\!\langle \underline{\mathsf{privenc}}^{\ell}(x_k, x_t) \rangle\!\rangle(\mathbf{I}, b) = \{\mathsf{symenc}(AV_k, AV_t, \ell, b) \mid AV_k \in \mathbf{I}(x_k), AV_t \in \mathbf{I}(x_t), AV_k = \mathsf{symkey}(\ldots)\} \cup \{\mathsf{symenc}(\mathsf{AnyPubVal}, AV_t, \ell, b) \mid \mathsf{AnyPubVal} \in \mathbf{I}(x_k), AV_t \in \mathbf{I}(x_t)\}$$

$$\langle\!\langle \underline{\mathsf{privdec}}(x_k, x_t) \rangle\!\rangle(\mathbf{I}, b) = \{AV_p \mid \mathsf{symenc}(AV_k, AV_p, \ell', b') \in \mathbf{I}(x_t), AV'_k \in \mathbf{I}(x_k), AV_k \cong_{\mathbf{P}} AV'_k\} \cup \{\mathsf{AnyPubVal} \mid \mathsf{AnyPubVal} \in \mathbf{I}(x_t), \mathbf{I}(x_k) \cap \mathbf{P} \neq \emptyset\} \cup \textit{if } \mathsf{AnyPubVal} \in \mathbf{I}(x_t) \textit{ then } \bigcup_{\mathsf{symkey}(i,\ell,b) \in \mathbf{I}(x_k)} \mathbf{E}_\ell^b \textit{ else } \emptyset$$

$$\langle\!\langle \underline{\mathsf{pubdec}}(x_k, x_t) \rangle\!\rangle(\mathbf{I}, b1) = \{AV_p \mid \mathsf{pubenc}(\mathsf{pubkey}(\ell'', b''), AV_p, \ell', b') \in \mathbf{I}(x_t), \mathsf{seckey}(\ell'', b'') \in \mathbf{I}(x_k)\} \cup \textit{if } \mathsf{AnyPubVal} \in \mathbf{I}(x_t) \textit{ then } \bigcup_{\mathsf{seckey}(\ell,b) \in \mathbf{I}(x_k)} \mathbf{E}_\ell^b \textit{ else } \emptyset$$

$$\langle\!\langle \underline{\mathsf{pubdec}}(x_k, x_t) \rangle\!\rangle(\mathbf{I}, b0) = \{\mathsf{AnyPubVal} \mid \mathsf{AnyPubVal} \in \mathbf{I}(x_t)\}$$

**Fig. 3.** Abstract semantics of expressions: mapping $\langle\!\langle e \rangle\!\rangle$

$$\langle\!\langle n \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \mathsf{false} \qquad \langle\!\langle x \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \mathsf{false}$$
$$\langle\!\langle \underline{\mathsf{keypair}}^{\ell} \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \mathsf{false} \qquad \langle\!\langle \mathsf{store}(x) \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \mathsf{false}$$

$$\langle\!\langle \mathsf{retrieve}(x) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \mathsf{AnyPubVal} \in \mathbf{I}(x) \vee \exists AV : \mathsf{store}(AV) \in \mathbf{I}(x)$$

$$\langle\!\langle \mathsf{retrieve}(x) \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \exists AV \in \mathbf{I}(x) : AV \neq \mathsf{store}(\ldots)$$
$$\langle\!\langle \mathsf{list}(x_1, \ldots, x_k) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \forall i : \exists AV \in \mathbf{I}(x_i) : AV \neq \mathsf{seckey}(\ldots)$$
$$\langle\!\langle \mathsf{list}(x_1, \ldots, x_k) \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \exists i : \exists AV \in \mathbf{I}(x_i) : AV = \mathsf{seckey}(\ldots)$$
$$\langle\!\langle \mathsf{gen\_symenc\_key}(i)^{\ell} \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \mathsf{false}$$
$$\langle\!\langle \underline{\pi_i^j}(x) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \mathsf{AnyPubVal} \in \mathbf{I}(x) \vee \exists (AV_i, \ldots, AV_j) \in \mathbf{I}(x)$$
$$\langle\!\langle \underline{\pi_i^j}(x) \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \exists AV \in \mathbf{I}(x) : AV \neq (AV_1, \ldots, AV_j)$$
$$\langle\!\langle \underline{\mathsf{pubkey}}(x) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \exists \ell, b : \mathsf{seckey}(\ell, b) \in \mathbf{I}(x)$$
$$\langle\!\langle \underline{\mathsf{pubkey}}(x) \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \exists AV \in \mathbf{I}(x) : AV \neq \mathsf{seckey}(\ldots)$$
$$\langle\!\langle \mathsf{gen\_nonce} \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) = \mathsf{false}$$

$$\langle\!\langle \underline{\mathsf{pubenc}}^{\ell}(x_k, x_t) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \mathsf{AnyPubVal} \in \mathbf{I}(x_k) \vee \exists \ell', b' : \mathsf{pubkey}(\ell', b') \in \mathbf{I}(x_k)$$

$$\langle\!\langle \underline{\mathsf{pubenc}}^{\ell}(x_{\mathrm{k}}, x_{\mathrm{t}}) \rangle\!\rangle_{\mathcal{E}}(\mathbf{I}, \mathbf{L}) = \{\mathsf{pubkey}(\ell', b') \in \mathbf{I}(x_{\mathrm{k}}) \wedge \mathbf{L} \Rightarrow \mathbf{I}(x_{\mathrm{t}}) \subseteq \mathbf{E}_{\ell'}^{b'} \mid \ell' \in \mathsf{L}_a\}$$

$$\langle\!\langle \underline{\mathsf{privenc}}^{\ell}(x_k, x_t) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \mathsf{AnyPubVal} \in \mathbf{I}(x_k) \vee \exists i, \ell', b' : \mathsf{symkey}(i, \ell', b') \in \mathbf{I}(x_k)$$

$$\langle\!\langle \underline{\mathsf{privenc}}^{\ell}(x_{\mathrm{k}}, x_{\mathrm{t}}) \rangle\!\rangle_{\mathcal{E}}(\mathbf{I}, \mathbf{L}) = \{\mathsf{symkey}(i, \ell', b') \in \mathbf{I}(x_{\mathrm{k}}) \wedge \mathbf{L} \Rightarrow \mathbf{I}(x_{\mathrm{t}}) \subseteq \mathbf{E}_{\ell'}^{b'} \mid \ell' \in \mathsf{L}_s\}$$

$$\langle\!\langle \underline{\mathsf{privdec}}(x_k, x_t) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \mathsf{AnyPubVal} \in \mathbf{I}(x_k) \vee \exists i, \ell', b' : \mathsf{seckey}(\ell', b') \in \mathbf{I}(x_k)$$
$$\langle\!\langle \underline{\mathsf{pubdec}}(x_k, x_t) \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) = \exists \ell', b' : \mathsf{seckey}(\ell', b') \in \mathbf{I}(x_k)$$

**Fig. 4.** Abstract semantics of expressions: mappings $\langle\!\langle e \rangle\!\rangle_{\mathrm{s}}$, $\langle\!\langle e \rangle\!\rangle_{\mathrm{f}}$ and $\langle\!\langle e \rangle\!\rangle_{\mathcal{E}}$

expression with respect to $\mathbf{X}$ that evaluates to a set of abstract values. The constraints in $\mathcal{C}$ relate the abstract values of variables that have been defined before the execution

$$\llbracket x := \mathsf{C}(x_1, \ldots, x_k) \rrbracket(b, \mathbf{X}, x) = \{C(AV_1, \ldots, AV_k) \mid AV_i \in \mathbf{X}(x_i)\}$$

$$\llbracket x := \mathsf{C}(x_1, \ldots, x_k) \rrbracket(b, \mathbf{X}, x_i) = \{AV_i \mid C(AV_1, \ldots, AV_k) \in \mathbf{X}(x)\}$$

$$\llbracket x := \underline{\pi_i^j}(y) \rrbracket(b, \mathbf{X}, x) = \{AV_i \mid (AV_1, \ldots, AV_k) \in \mathbf{X}(y)\} \cup \{\mathsf{AnyPubVal} \mid \mathsf{AnyPubVal} \in \mathbf{X}(y)\}$$

$$\begin{aligned}\llbracket x := \underline{\pi_i^j}(y) \rrbracket(b, \mathbf{X}, y) = \; & \{(AV'_1, \ldots, AV'_{i-1}, AV_i \sqcap AV'_i, AV'_{i+1}, \ldots, AV'_j) \mid \\ & AV_i \in \mathbf{X}(x), (AV'_1, \ldots, AV'_j) \in \mathbf{X}(y)\} \cup \\ & \{(\mathsf{AnyPubVal}^{i-1}, AV_i, \mathsf{AnyPubVal}^{j-i}) \mid \\ & AV_i \in \mathbf{X}(x), \mathsf{AnyPubVal} \in \mathbf{X}(y)\}\end{aligned}$$

$$\llbracket x := y \rrbracket(b, \mathbf{X}, x) = \mathbf{X}(y)$$
$$\llbracket x := y \rrbracket(b, \mathbf{X}, y) = \mathbf{X}(x)$$

$$\llbracket y := \underline{\mathsf{pubkey}}(x) \rrbracket(b, \mathbf{X}, y) = \{\mathsf{pubkey}(\ell, b') \mid \mathsf{seckey}(\ell, b') \in \mathbf{X}(x)\}$$

$$\llbracket y := \underline{\mathsf{pubkey}}(x) \rrbracket(b, \mathbf{X}, x) = \{\mathsf{seckey}(\ell, b') \mid \mathsf{pubkey}(\ell, b') \in \mathbf{X}(x)\}$$

$$\llbracket y := \underline{\mathsf{pubdec}}(x_k, x_t) \rrbracket(b1, \mathbf{X}, y) = \mathit{if}\ \mathsf{AnyPubVal} \in \mathbf{X}(x_k)\ \mathit{then}\ \mathbf{X}(y)\ \mathit{else} \bigcup_{\mathsf{seckey}(\ell', b') \in \mathbf{X}(x_k)} \mathbf{E}_{\ell'}^{b'}$$

$$\llbracket y := \underline{\mathsf{privdec}}(x_k, x_t) \rrbracket(b, \mathbf{X}, y) = \mathit{if}\ \mathsf{AnyPubVal} \in \mathbf{X}(x_k)\ \mathit{then}\ \mathbf{X}(y)\ \mathit{else} \bigcup_{\mathsf{symkey}(i, \ell', b') \in \mathbf{X}(x_k)} \mathbf{E}_{\ell'}^{b'}$$

**Fig. 5.** Constraints giving the relationships between values of variables

of $P$. If $\mathcal{C}$ contains a constraint $\mathbf{X}(x) \dot{\subseteq} E$ then the result of $E$ is a suitable abstraction for the value of $x$.

Let $\mathbf{R}$ be a mapping from variables to sets of abstract values, and let $\mathcal{C}$ be a set of constraints in the form $\mathbf{X}(x) \dot{\subseteq} E$. We let $\mathfrak{L}(\mathbf{R}, \mathcal{C})$ denote the greatest solution to the constraints $\mathcal{C}$ (with $\mathbf{X}$ as the variable) that is less than or equal to $\mathbf{R}$. The mapping $\langle\!\langle\!\langle P \rangle\!\rangle\!\rangle$ is given in Fig. 6.

Let us explain the generated constraints. For an assignment, the following constraints are generated. If the process is alive ($\mathbf{L}$ is $\mathsf{true}$) and the expression $e$ may succeed [resp. fail] then we demand that the boolean variable reflecting that — $\mathbf{L}^b_{\ell,\mathsf{true}}$ [resp. $\mathbf{L}^b_{\ell,\mathsf{false}}$] is $\mathsf{true}$, too. If $e$ may succeed and hence $\mathbf{L}^b_{\ell,\mathsf{true}}$ is $\mathsf{true}$ then we let the mapping $\mathbf{R}^b_\ell$ be (at least) the mapping $\mathbf{I}$, but additionally we fix the abstraction of the left-hand side $y$. Here this "at least" means "equal to" because there will be no other constraints for $\mathbf{R}^b_\ell$. If $\mathbf{L}^b_{\ell,\mathsf{true}}$ is $\mathsf{false}$ then $\mathbf{R}^b_\ell$ has no constraints, hence it maps everything to $\emptyset$.

The set of constraints $\mathcal{C}'\langle b \rangle$ includes all the relationships between variables that are defined up to the successful execution of $y := e$. Note that the inequality signs in the constraints in $\mathcal{C}'\langle b \rangle$ has the opposite direction from the inequality signs in the constraints generated by $\langle\!\langle\!\langle P \rangle\!\rangle\!\rangle$. The constraint $\mathbf{S}^b_\ell \geq \mathfrak{L}(\mathbf{R}^b_\ell, \mathcal{C}'\langle b \rangle)$ states that $\mathbf{S}^b_\ell$ contains basically the same abstractions as $\mathbf{R}^b_\ell$, but all recorded relationships between variables have been taken into account. As this constraint is the only one for $\mathbf{S}^b_\ell$, the inequality $\mathbf{S}^b_\ell \leq \mathbf{R}^b_\ell$ always holds.

We also add the constraints for the variables $\mathbf{E}^{b'}_{\ell'}$ and we recursively invoke $\langle\!\langle\!\langle \cdot \rangle\!\rangle\!\rangle$ for the default- and the false-branch. The arguments for these recursive calls are also worth noting. We see that as we pass through the protocol, we collect the constraints expressing the relationships between variables. For the default-branch, $\mathbf{S}^b_\ell$ is the abstraction of the initial state, while for the false-branch, the same mapping $\mathbf{I}$ is used because no variable was assigned to if the evaluation of $e$ failed. Also we collect no new relationships between variables if $e$ fails (although it would be possible for some $e$, we have found that it does not change the precision of the analysis in practice).

If $e$ is a public-key decryption then we have "two different default-branches", with decryption contexts $b1$ and $b0$. These two default-branches are reflected in the variables $\mathbf{R}^{b1}_\ell$, $\mathbf{S}^{b1}_\ell$, $\mathbf{R}^{b0}_\ell$, $\mathbf{S}^{b0}_\ell$, as well as in two invocations of $\langle\!\langle\!\langle \cdot \rangle\!\rangle\!\rangle$ for the default-branch $P$.

Consider now other cases for the process $P$. In an if-statement we check whether the abstractions of $x$ and $x'$ may intersect. We also add the equality of $x$ and $x'$ to the set of constraints $\mathcal{C}'$.

A send-command always succeeds, the intended recipient becomes known to the adversary and the message is recorded as occurring on the channel $c$ and/or becoming known to the adversary.

In a receive-statement (no matter whether replicated or not), a message from the adversary is abstracted as $\mathsf{AnyPubVal}$ and a message from the user as a secret payload. The sender of the message is already known to the adversary, hence $x_{\mathrm{p}}$ is a public integer. No new constraints are added, hence the invocation of $\mathfrak{L}$ is not needed for $\mathbf{S}^b_\ell$.

If $I_i$ is the program for the machine $\mathsf{P}_i$ then the set of constraints for the protocol is the union of $\langle\!\langle\!\langle I_i \rangle\!\rangle\!\rangle(\{\}, \varepsilon, \mathsf{true}, \emptyset)$ over all $i$, where $\{\}$ is the mapping with empty domain. Together with the constraints in Fig. 1, they make up the entire constraint system that we have to solve.

Suppose that we have solved the constraint system for some protocol $\wp$. Let $\ell$ be a label occurring in this protocol (labeling a subprocess $P$). Let $\mathbf{I}^b_\ell$ denote the variable $\mathbf{S}^{b'}_{\ell'}$

If $e$ is a not a public-key decryption then

$$\begin{aligned}
&\langle\!\langle\!\langle \mathbf{let}^{\ell}\ y := e\ \mathbf{in}\ P\ \mathbf{else}\ P' \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \\
&\qquad let\ \mathcal{C}'\langle b\rangle = \mathcal{C} \cup \{\mathbf{X}(y) \dot{\subseteq} \lfloor\!\lfloor e \rfloor\!\rfloor(b, \mathbf{X}, y) \mid y \in \mathbf{Var}^{\bullet}_{\ell}\}\ in \\
&\qquad \{\mathbf{L} \wedge \langle\!\langle e \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) \Rightarrow \mathbf{L}^{b}_{\ell,\mathsf{true}},\ \mathbf{L} \wedge \langle\!\langle e \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) \Rightarrow \mathbf{L}^{b}_{\ell,\mathsf{false}}, \\
&\qquad \mathbf{L}^{b}_{\ell,\mathsf{true}} \Rightarrow \mathbf{R}^{b}_{\ell} \geq \mathbf{I}[y \mapsto \langle\!\langle e \rangle\!\rangle(\mathbf{I}, b)],\ \mathbf{S}^{b}_{\ell} \geq \mathfrak{L}(\mathbf{R}^{b}_{\ell}, \mathcal{C}'\langle b\rangle)\} \cup \\
&\qquad \langle\!\langle e \rangle\!\rangle_{\mathcal{E}}(\mathbf{I}, \mathbf{L}^{b}_{\ell,\mathsf{true}}) \cup \langle\!\langle\!\langle P \rangle\!\rangle\!\rangle(\mathbf{S}^{b}_{\ell}, b, \mathbf{L}^{b}_{\ell,\mathsf{true}}, \mathcal{C}'\langle b\rangle) \cup \\
&\qquad\qquad \langle\!\langle\!\langle P' \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}^{b}_{\ell,\mathsf{false}}, \mathcal{C})
\end{aligned}$$

If $e$ is a public-key decryption then

$$\begin{aligned}
&\langle\!\langle\!\langle \mathbf{let}^{\ell}\ y := e\ \mathbf{in}\ P\ \mathbf{else}\ P' \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \\
&\qquad let\ \mathcal{C}'\langle b\rangle = \mathcal{C} \cup \{\mathbf{X}(y) \dot{\subseteq} \lfloor\!\lfloor e \rfloor\!\rfloor(b, \mathbf{X}, y) \mid y \in \mathbf{Var}^{\bullet}_{\ell}\}\ in \\
&\qquad \{\mathbf{L} \wedge \langle\!\langle e \rangle\!\rangle_{\mathrm{s}}(\mathbf{I}) \Rightarrow \mathbf{L}^{b}_{\ell,\mathsf{true}},\ \mathbf{L} \wedge \langle\!\langle e \rangle\!\rangle_{\mathrm{f}}(\mathbf{I}) \Rightarrow \mathbf{L}^{b}_{\ell,\mathsf{false}}, \\
&\qquad \mathbf{L}^{b}_{\ell,\mathsf{true}} \Rightarrow \mathbf{R}^{b1}_{\ell} \geq \mathbf{I}[y \mapsto \langle\!\langle e \rangle\!\rangle(\mathbf{I}, b1)],\ \mathbf{S}^{b1}_{\ell} \geq \mathfrak{L}(\mathbf{R}^{b1}_{\ell}, \mathcal{C}'\langle b1\rangle), \\
&\qquad \mathbf{L}^{b}_{\ell,\mathsf{true}} \Rightarrow \mathbf{R}^{b0}_{\ell} \geq \mathbf{I}[y \mapsto \langle\!\langle e \rangle\!\rangle(\mathbf{I}, b0)],\ \mathbf{S}^{b0}_{\ell} \geq \mathfrak{L}(\mathbf{R}^{b0}_{\ell}, \mathcal{C}'\langle b0\rangle)\} \cup \\
&\qquad \langle\!\langle e \rangle\!\rangle_{\mathcal{E}}(\mathbf{I}, \mathbf{L}^{b}_{\ell,\mathsf{true}}) \cup \langle\!\langle\!\langle P \rangle\!\rangle\!\rangle(\mathbf{S}^{b1}_{\ell}, b1, \mathbf{L}^{b}_{\ell,\mathsf{true}}, \mathcal{C}'\langle b1\rangle) \cup \\
&\qquad\qquad \langle\!\langle\!\langle P \rangle\!\rangle\!\rangle(\mathbf{S}^{b0}_{\ell}, b0, \mathbf{L}^{b}_{\ell,\mathsf{true}}, \mathcal{C}'\langle b0\rangle) \cup \langle\!\langle\!\langle P' \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}^{b}_{\ell,\mathsf{false}}, \mathcal{C})
\end{aligned}$$

$$\begin{aligned}
&\langle\!\langle\!\langle \mathbf{if}^{\ell}\ x = x'\ \mathbf{then}\ P\ \mathbf{else}\ P' \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \\
&\qquad let\ \mathcal{C}' = \mathcal{C} \cup \{\mathbf{X}(x) \dot{\subseteq} \mathbf{X}(x'), \mathbf{X}(x') \dot{\subseteq} \mathbf{X}(x)\}\ in \\
&\qquad \{\mathbf{L} \wedge (\exists AV \in \mathbf{I}(x)\, \exists AV' \in \mathbf{I}(x') : AV \cong_{\mathbf{P}} AV') \Rightarrow \mathbf{L}^{b}_{\ell,\mathsf{true}}, \\
&\qquad\qquad \mathbf{L} \Rightarrow \mathbf{L}^{b}_{\ell,\mathsf{false}},\ \mathbf{R}^{b}_{\ell} \geq \mathbf{I},\ \mathbf{S}^{b}_{\ell} \geq \mathfrak{L}(\mathbf{R}^{b}_{\ell}, \mathcal{C}')\} \cup \\
&\qquad\qquad\qquad \langle\!\langle\!\langle P \rangle\!\rangle\!\rangle(\mathbf{S}^{b}_{\ell}, b, \mathbf{L}^{b}_{\ell,\mathsf{true}}, \mathcal{C}') \cup \langle\!\langle\!\langle P' \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}^{b}_{\ell,\mathsf{false}}, \mathcal{C})
\end{aligned}$$

$$\begin{aligned}
&\langle\!\langle\!\langle \mathbf{send}_{c}[x_{\mathrm{p}}](x).I^{*} \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \langle\!\langle\!\langle I^{*} \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) \cup \\
&\qquad\qquad \{\mathbf{L} \Rightarrow \mathbf{I}(x_{\mathrm{p}}) \subseteq \mathbf{P}, \\
&\qquad\qquad \mathbf{L} \wedge (c \in \mathbf{Chan}_{\mathsf{s}} \cup \mathbf{Chan}_{\mathsf{a}}) \Rightarrow \mathbf{I}(x) \subseteq \mathbf{C}_{c}, \\
&\qquad\qquad\qquad \mathbf{L} \wedge (c \in \mathbf{Chan}_{\mathsf{a}} \cup \mathbf{Chan}_{\mathsf{i}}) \Rightarrow \mathbf{I}(x) \subseteq \mathbf{P}\}
\end{aligned}$$

$$\begin{aligned}
&\langle\!\langle\!\langle (!)\mathbf{receive}^{\ell}_{c}[x_{\mathrm{p}}](x).P \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \\
&\qquad \{\mathbf{L} \Rightarrow \mathbf{L}^{\ell}_{b,\mathsf{true}},\ \mathbf{S}^{b}_{\ell} \geq \mathbf{R}^{b}_{\ell}\} \cup \langle\!\langle\!\langle P \rangle\!\rangle\!\rangle(\mathbf{S}^{b}_{\ell}, b, \mathbf{L}^{b}_{\ell,\mathsf{true}}, \mathcal{C}) \cup \\
&\qquad\qquad \begin{cases}
\{\mathbf{L} \Rightarrow \mathbf{R}^{b}_{\ell} \geq \mathbf{I}[x \mapsto \mathbf{C}_{c}, x_{\mathrm{p}} \mapsto \{\mathsf{X_P}\}]\}, \\
\qquad\qquad \text{if } c \in \mathbf{Chan}_{\mathsf{s}} \cup \mathbf{Chan}_{\mathsf{a}} \\
\{\mathbf{L} \Rightarrow \mathbf{R}^{b}_{\ell} \geq \mathbf{I}[x \mapsto \{\mathsf{AnyPubVal}\}, x_{\mathrm{p}} \mapsto \{\mathsf{X_P}\}]\}, \\
\qquad\qquad \text{if } c \in \mathbf{Chan}_{\mathsf{i}} \\
\{\mathbf{L} \Rightarrow \mathbf{R}^{b}_{\ell} \geq \mathbf{I}[x \mapsto \{\mathsf{X_S}\}, x_{\mathrm{p}} \mapsto \{\mathsf{X_P}\}]\}, \\
\qquad\qquad \text{if } c \in \mathbf{Chan}_{\mathsf{u}}
\end{cases}
\end{aligned}$$

$$\langle\!\langle\!\langle I_1 \mid \cdots \mid I_n \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \bigcup_{i=1}^{n} \langle\!\langle\!\langle I_i \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C})$$

$$\langle\!\langle\!\langle \mathbf{0} \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \langle\!\langle\!\langle \mathtt{JJ} \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C}) = \emptyset\ .$$

**Fig. 6.** Constraints generated by processes

(or the empty mapping) that occurs as the first argument in the call $\langle\!\langle\!\langle P^\ell \rangle\!\rangle\!\rangle(\mathbf{I}, b, \mathbf{L}, \mathcal{C})$, invoked during the construction of constraints for $\wp$. That is, $\mathbf{I}^b_\ell$ gives the abstract values of variables before entering the subprocess labeled with $\ell$ in the context $b$.

The following theorem states how the security of a protocol can be established using our analysis.

**Theorem 1.** *Let $\wp$ be a protocol and let Let $\mathbf{S}^b_\ell$, $\mathbf{E}^b_\ell$, $\mathbf{P}$, $\mathbf{C}_c$, $\mathbf{L}^b_{\ell,\mathfrak{b}}$ be such that the constraints given above are fulfilled. If the following conditions hold then the composition of machines $\mathcal{TH}_n$ and $\mathsf{P}_i$ ($1 \leq i \leq n$) preserves the secrecy of payloads, i.e., the payloads are cryptographically secret if $\mathcal{TH}_n$ is replaced by its cryptographic realization.*

*(I)* *If the protocol contains a statement of the form $\mathbf{if}^\ell\ x = x' \ldots$ then $\mathsf{X_S} \notin \mathbf{I}^b_\ell(x)$, $\mathsf{X_S} \notin \mathbf{I}^b_\ell(x')$, $\mathsf{store}(\mathsf{X_S}) \notin \mathbf{I}^b_\ell(x)$ and $\mathsf{store}(\mathsf{X_S}) \notin \mathbf{I}^b_\ell(x')$ for any $b$.*

*(II)* *If $\mathsf{X_S} \in \mathbf{S}^b_\ell(x)$ for some $b$, $x$, and this $x$ occurs as an argument to some operation where the abstract values at entry are given by $\mathbf{S}^b_\ell$, then this operation is* store *or a* send *to a user.*

*(III)* $\mathsf{X_S} \notin \mathbf{P}$.

*(IV)* *If $\mathsf{AV} \in \mathbf{E}^b_\ell$ and a symm. key of order $i$ is generated at $\ell$ then the order of $\mathsf{AV}$ is less than $i$.*

- *The order of $\mathsf{symkey}(i, \ell, b)$ is $i$. The order of a tuple is the maximum order of its members. The order of other abstract values is $0$.*

*(V)* $\mathsf{symkey}(i, \ell, b) \notin \mathbf{P}$ *for any $i, \ell, b$.*

## 4 Implementation

We find the (componentwise) least solution for the aforementioned collection of inequalities. The least fixed point is computed iteratively, using a version of the solver from [65], which is specifically tailored to systems of constraints. The computation might not terminate but we believe that this is not a problem for real protocols; in fact, we have never encountered this situation when we applied our tool to common protocols of the literature. The only case in which computation is not guaranteed to terminate is if the protocol is able to create terms of arbitrary complexity all by itself, without the help from the adversary. We also believe that the potentially exponential number of variables in the size of the protocol is not a problem in practice because protocol descriptions are short.

There are ways to deal with the divergence of the fixed-point computation (add suitable widenings). Also, we believe that the exponential number of sets can be represented in a more compact way, if necessary.

The value of $\mathfrak{L}(\mathbf{R}, \mathcal{C})$ is also computed iteratively, using the same solution method. The mapping $\mathbf{X}$ is initialized with $\mathbf{R}$ and the iteration proceeds downwards.

The constraint generator and solver have been implemented in O'Caml (version 3.09 was used to compile it to native code). We have tested the analyser on several protocols from the literature, namely Needham-Schroeder public key [66], its fix by Lowe [8], Otway-Rees [67], Yahalom, and its modification by Burrows et. al [68]. The goal of all these protocols is to exchange a symmetric key between two parties. We use our analysis to find out whether it is safe to use the exchanged key to protect secret

payloads in transit over public networks. We have thus added an extra message at the end of the protocol sessions in all of these protocols. This extra message contains a secret payload encrypted under the freshly exchanged key. (Note again that we do not consider real-or-random secrecy of the keys in this paper, but secrecy of the (encrypted) payloads.) Our analysis considers all these protocols secure, except for the original (and indeed flawed) Needham-Schroeder protocol. If one allows old session keys to become known to the adversary, there may be attacks that are not discovered by our analyser since the BPW model does not consider the leakage of secret keys that have been used (this would cause a so-called commitment problem which makes a proof of computational soundness in the sense of BRSIM/UC impossible). Problems with the Needham-Schroeder-Lowe and the modified Yahalom protocol have been published [42, 69] but these problems do not affect payload secrecy properties — they do not allow an adversary to learn the new key or inject its own key which could then be used to learn information on the payload. The running times of the analyser on a computer with 1 GHz Intel Celeron processor and 256 MB of main memory are between one and eight seconds for these protocols with less than two seconds for Needham-Schroeder-Lowe and both versions of Yahalom.

## 5 Correctness of the Analysis

Theorem 1 is a straightforward corollary of a lemma similar to subject reduction. We are going to give the statement of that lemma here, and for the sake of readability postpone the technical proof to [70].

When arguing about the correctness, we need to distinguish public and secret payloads. Hence we change the semantics of the system a little bit and assume that each payload is labeled with its secrecy level. An integer received from the protocol user is labeled as secret; a constant integer or the apparent sender of some message is labeled as public. When the payloads are stored in the database of terms then the labels are stored as well. They are also retrieved together with labels. When two integers are compared their secrecy levels are ignored.

Let $\mathfrak{R}$ be a run of the protocol (a finite sequence of steps). Let $\mathcal{O}$ be the state of the database of $\mathcal{TH}_n$ at the end of this run. Let $x$ be a protocol variable whose definition occurs under $k$ replications. If the variable $x$ has been assigned a value inside the $i_1$-th replica of the outermost replication, $i_2$-th replica of the next replication, etc., then we denote this value $\mathcal{O}(x, [i_1, \ldots, i_k])$. This value is either a handle to a term in $\mathcal{O}$, a secret integer, or a public integer. Similarly, if a program point $\ell$ inside $k$ replications and (syntactically) preceded by $n$ public-key decryptions then let $\mathcal{O}(\ell, [i_1, \ldots, i_k])$ be the $n$-bit string whose bits describe the source of $n$ ciphertexts whose decryptions are visible at the point $\ell$ in the replica $[i_1, \ldots, i_k]$. Note that a ciphertext is a term in the database $\mathcal{O}$ and its creator is simply the first principal that had a handle to it.

For a set $\mathcal{T}$ of terms in the database $\mathcal{O}$ we let its *downwards closure* $\overline{\mathcal{T}}$ be the smallest set of terms containing $\mathcal{T}$ and being closed with respect to list projection, decryption with keys in $\overline{\mathcal{T}}$, and extracting the public keys and symmetric key names from ciphertexts. For an abstract value $AV$ we define its semantics $[\![AV]\!]_{\mathcal{O}}$ with respect to the

contents of the database $\mathcal{O}$. The semantics is either a set of terms in $\mathcal{O}$ or a set of payloads.

- $[\![\mathsf{X_P}]\!]_{\mathcal{O}} = \{\text{“public } n\text{”} \mid n \in \mathbb{N}\}$.
- $[\![\mathsf{X_S}]\!]_{\mathcal{O}} = \{\text{“secret } n\text{”} \mid n \in \mathbb{N}\}$.
- $[\![\mathsf{store}(AV)]\!]_{\mathcal{O}}$ is the set of all terms of type data in $\mathcal{O}$ whose argument belongs to $[\![AV]\!]_{\mathcal{O}}$.
- $[\![\mathsf{nonce}(\ell, b)]\!]_{\mathcal{O}}$ is the set of all terms of type nonce in $\mathcal{O}$ that are generated by the gen_nonce-expressions at the protocol point $\ell$ at replicas $[i_1, \ldots, i_k]$, such that $b = \mathcal{O}(\ell, [i_1, \ldots, i_k])$
- $[\![\mathsf{symkey}(i, \ell, b)]\!]_{\mathcal{O}}$, $[\![\mathsf{symkeyname}(\ell, \mathsf{b})]\!]_{\mathcal{O}}$, $[\![\mathsf{seckey}(\ell, b)]\!]_{\mathcal{O}}$, and $[\![\mathsf{pubkey}(\ell, b)]\!]_{\mathcal{O}}$ — defined the same way as $[\![\mathsf{nonce}(\ell, b)]\!]_{\mathcal{O}}$ (only replace nonce with skse, pkse, ske, or pke).
- $[\![(AV_1, \ldots, AV_j)]\!]_{\mathcal{O}}$ is the set of all terms of type list in $\mathcal{O}$ whose length is $j$ and whose $i$-th component term ($1 \leq i \leq j$) belongs to $[\![AV_i]\!]_{\mathcal{O}}$.
- $[\![\mathsf{pubenc}(AV_k, AV_p, \ell, b)]\!]_{\mathcal{O}}$ is the set of all terms of type enc, such that
  - they have been created by the pubenc-expressions at the protocol point $\ell$ at replicas $[i_1, \ldots, i_k]$, such that $b = \mathcal{O}(\ell, [i_1, \ldots, i_k])$;
  - the term corresponding to the public key must belong to $[\![AV_k]\!]_{\mathcal{O}}$;
  - the term corresponding to the plaintext must belong to $[\![AV_p]\!]_{\mathcal{O}}$.
- $[\![\mathsf{symenc}(AV_k, AV_p, \ell, b)]\!]_{\mathcal{O}}$ is defined similarly, where enc is replaced with symenc.
- $[\![\mathsf{AnyPubVal}]\!]_{\mathcal{O}}$ is the downwards closure of the set of all terms that the adversary knows.

Let $\tilde{\mathbf{P}}$ be the largest set that $\tilde{\mathbf{P}} \subseteq \mathbf{P} \backslash \{\mathsf{AnyPubVal}\}$ and $(AV_1, \ldots, AV_j) \in \tilde{\mathbf{P}}$ implies $AV_i \in \tilde{\mathbf{P}}$ for $1 \leq i \leq j$. Informally, $\tilde{\mathbf{P}}$ is obtained from $\mathbf{P}$ by deleting AnyPubVal and also any abstract value that is a list, one of whose components (after flattening lists) is AnyPubVal. The set $\tilde{\mathbf{P}}$ is a better characterization than $\mathbf{P}$ for the set of terms created by honest participants and learned by the adversary.

**Definition 1.** *A term $T$ from the downwards closure of the set of all terms known to the adversary is* adversarially well-constructed *with respect to $\mathbf{P}$ if one of the following holds:*

- $\exists AV \in \tilde{\mathbf{P}}$*, such that* $T \in [\![AV]\!]_{\mathcal{O}}$.
- *All immediate subterms of $T$ (the immediate subterm of a public key or a symmetric key name is the corresponding secret key) are known to the adversary and are also adversarially well-constructed. Also, if $T$ is of type* nonce, ske, enc, garbage, skse, symenc *then $T$ must have been constructed by the adversary.*

That is, a term is adversarially well-constructed if the adversary knows how to construct this term from the terms that the analysis has found him to know.

**Lemma 1 (Subject reduction).** *Let $\wp$ be a protocol and let Let $\mathbf{S}^b_\ell$, $\mathbf{E}^b_\ell$, $\mathbf{P}$, $\mathbf{C}_c$, $\mathbf{L}^b_{\ell,\mathsf{b}}$ be such that the constraints given in Sec. 3.3 are fulfilled. Let $\mathfrak{R}$ be a run of the protocol $\wp$. Let $\mathcal{O}$ be the state of the database of $\mathcal{TH}_n$ at the end of $\mathfrak{R}$. The following claims hold.*

$\boxed{P}$ *If a term $T$ is known to adversary then it is adversarially well-constructed wrt. $\mathbf{P}$.*

$\boxed{X}$ *If* $\mathcal{O}(x, [i_1, \ldots, i_k]) = T$ *(here* $T$ *may be both a term or an immediate value) and the replica* $[i_1, \ldots, i_k]$ *passes through the point* $\ell$ *with the operation at* $\ell$ *succeeding and the value of* $x$ *being defined, then there exists* $AV \in \mathbf{S}_\ell^{\mathcal{O}(\ell,[i_1,\ldots,i_k])}(x)$*, such that* $T \in [\![AV]\!]_{\mathcal{O}}$*.*

$\boxed{C}$ *If a term* $T$ *is communicated over an abstract channel* $c \in \mathbf{Chan_s} \cup \mathbf{Chan_a}$ *then there exists* $AV \in \mathbf{C}_c$*, such that* $T \in [\![AV]\!]_{\mathcal{O}}$*.*

$\boxed{E}$ *If* $T_k$ *is the term representing an asymmetric or symmetric key generated at the program point* $\ell$ *in the replica* $[i_1, \ldots, i_k]$*) and* $T_p$ *is a term that occurs as the plaintext in an encryption where* $T_k$ *is the key, then at least one of the following holds:*

- *there exists* $AV \in \mathbf{E}_\ell^{\mathcal{O}(\ell,[i_1,\ldots,i_k])}$*, such that* $T_p \in [\![AV]\!]_{\mathcal{O}}$*.*
- $T_k$ *and* $T_p$ *are both known to the adversary and the term representing encryption of* $T_p$ *with* $T_k$ *is constructed by the adversary.*

$\boxed{L}$ *If* $\ell$ *is a branching point in the protocol (a* $\mathbf{let}$*- or* $\mathbf{if}$*-statement) and if the* $B$*-branch was taken at the replica* $[i_1, \ldots, i_k]$ *(here* $B$ *is either true/default or false), then* $\mathbf{L}_{\ell,B}^{\mathcal{O}(\ell,[i_1,\ldots,i_k])} = \mathsf{true}$*.*

The lemma is proved by induction over the length of $\mathfrak{R}$.

## 6 Acknowledgments

We thank the anonymous reviewers for their helpful comments. M. Backes was supported by the German Research Foundation (DFG) under grant 3194/1-1. P. Laud was supported by Estonian Science Foundation, grant #6095, and by EU Integrated Project AEOLUS (contract no. IST-15964). The constraint solver used in the implementation is from the project Goblin [71].

## References

1. Michael Backes and Peeter Laud. Computationally sound secrecy proofs by mechanized flow analysis. In *Proceedings of 13th ACM Conference on Computer and Communications Security (CCS)*, pages 370–379, 2006.
2. Danny Dolev and Andrew C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–208, 1983.
3. Shimon Even and Oded Goldreich. On the security of multi-party ping-pong protocols. In *Proc. 24th IEEE FOCS*, pages 34–39, 1983.
4. Michael Merritt. *Cryptographic Protocols*. PhD thesis, Georgia Institute of Technology, 1983.
5. Richard Kemmerer, Catherine Meadows, and Jon Millen. Three systems for cryptographic protocol analysis. *Journal of Cryptology*, 7(2):79–130, 1994.
6. Steve Schneider. Security properties and CSP. In *Proc. 17th IEEE Symp. on Security & Privacy*, pages 174–187, 1996.
7. M. Abadi and A. D. Gordon. A calculus for cryptographic protocols: The spi calculus. In *Proc. 4th ACM CCS*, pages 36–47, 1997.
8. Gavin Lowe. Breaking and fixing the Needham-Schroeder public-key protocol using FDR. In *Proc. 2nd TACAS*, volume 1055 of *LNCS*, pages 147–166. Springer, 1996.

9. Lawrence Paulson. The inductive approach to verifying cryptographic protocols. *Journal of Cryptology*, 6(1):85–128, 1998.
10. David Basin, Sebastian Mödersheim, and Luca Viganò. OFMC: A symbolic model checker for security protocols. *Intern. Journal of Information Security*, 2004.
11. Peeter Laud. Secrecy types for a simulatable cryptographic library. In *Proc. 12th ACM CCS*, pages 26–35, 2005.
12. Birgit Pfitzmann and Michael Waidner. Composition and integrity preservation of secure reactive systems. In *Proc. 7th ACM CCS*, pages 245–254, 2000.
13. Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *Proc. 42nd IEEE FOCS*, pages 136–145, 2001.
14. Birgit Pfitzmann and Michael Waidner. A model for asynchronous reactive systems and its application to secure message transmission. In *Proc. 22nd IEEE Symp. on Security & Privacy*, pages 184–200, 2001. Extended version (with M. Backes) in IACR ePrint Report 2004/082.
15. Michael Backes, Birgit Pfitzmann, and Michael Waidner. The reactive simulatability framework for asynchronous systems. *Information and Computation*, pages 1685–1720, 2007.
16. Michael Backes and Birgit Pfitzmann. A cryptographically sound security proof of the Needham-Schroeder-Lowe public-key protocol. In *Proc. 23rd Conf. on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 1–12, 2003.
17. Michael Backes. A cryptographically sound dolev-yao style security proof of the Otway-Rees protocol. In *Proceedings of 9th European Symposium on Research in Computer Security (ESORICS)*, volume 3193 of *Lecture Notes in Computer Science*, pages 89–108. Springer, 2004.
18. Michael Backes and Markus Duermuth. A cryptographically sound Dolev-Yao style security proof of an electronic payment system. In *Proceedings of 18th IEEE Computer Security Foundations Workshop (CSFW)*, pages 78–93, 2005.
19. Michael Backes and Birgit Pfitzmann. On the cryptographic key secrecy of the strengthened Yahalom protocol. In *Proceedings of 21st IFIP International Information Security Conference (SEC)*, pages 233–245, 2006.
20. Michael Backes, Sebastian Moedersheim, Birgit Pfitzmann, and Luca Vigano. Symbolic and cryptographic analysis of the secure WS-ReliableMessaging Scenario. In *Proceedings of Foundations of Software Science and Computational Structures (FOSSACS)*, volume 3921 of *Lecture Notes in Computer Science*, pages 428–445. Springer, 2006.
21. Michael Backes, Iliano Cervesato, Aaron D. Jaggard, Andre Scedrov, and Joe-Kai Tsay. Cryptographically sound security proofs for basic and public-key kerberos. In *Proceedings of 11th European Symposium on Research in Computer Security(ESORICS)*, volume 4189 of *Lecture Notes in Computer Science*, pages 362–383. Springer, 2006. Preprint on IACR ePrint 2006/219.
22. Christoph Sprenger, Michael Backes, David Basin, Birgit Pfitzmann, and Michael Waidner. Cryptographically sound theorem proving. In *Proc. 19th IEEE CSFW*, 2006.
23. Martín Abadi and Bruno Blanchet. Secrecy types for asymmetric communication. *Theoretical Computer Science*, 298(3):387–415, 2003.
24. Reinhard Wilhelm, Shmuel Sagiv, and Thomas W. Reps. Shape Analysis. In *Proc. 9th CC*, volume 1781 of *LNCS*, pages 1–17. Springer, 2000.
25. Martin Abadi and Phillip Rogaway. Reconciling two views of cryptography: The computational soundness of formal encryption. In *Proc. 1st IFIP TCS*, volume 1872 of *LNCS*, pages 3–22. Springer, 2000.
26. J. D. Guttman, F. J. Thayer Fabrega, and L. Zuck. The faithfulness of abstract protocol analysis: Message authentication. In *Proc. 8th ACM CCS*, pages 186–195, 2001.
27. Peeter Laud. Semantics and program analysis of computationally secure information flow. In *Proc. 10th ESOP*, pages 77–91, 2001.

28. Jonathan Herzog, Moses Liskov, and Silvio Micali. Plaintext awareness via key registration. In *Proc. CRYPTO 2003*, volume 2729 of *LNCS*, pages 548–564. Springer, 2003.
29. Michael Backes and Birgit Pfitzmann. Relating symbolic and cryptographic secrecy. *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 2(2):109–123, 2005.
30. Michael Backes, Birgit Pfitzmann, and Michael Waidner. A composable cryptographic library with nested operations. In *Proc. 10th ACM CCS*, pages 220–230, 2003.
31. Michael Backes, Birgit Pfitzmann, and Michael Waidner. Symmetric authentication within a simulatable cryptographic library. In *Proc. 8th ESORICS*, volume 2808 of *LNCS*, pages 271–290. Springer, 2003.
32. Michael Backes and Birgit Pfitzmann. Symmetric encryption in a simulatable Dolev-Yao style cryptographic library. In *Proc. 17th IEEE CSFW*, pages 204–218, 2004.
33. Michael Backes, Birgit Pfitzmann, and Andre Scedrov. Key-dependent message security under active attacks - BRSIM/UC-soundness of symbolic encryption with key cycles. In *Proceedings of 20th IEEE Computer Security Foundation Symposium (CSF)*, 2007. Preprint on IACR ePrint 2005/421.
34. Michael Backes and Birgit Pfitzmann. Computational probabilistic non-interference. In *Proceedings of 7th European Symposium on Research in Computer Security (ESORICS)*, volume 2502 of *Lecture Notes in Computer Science*, pages 1–23. Springer, 2002.
35. Michael Backes and Christian Jacobi. Cryptographically sound and machine-assisted verification of security protocols. In *Proc. 20th Annual Symp. on Theoretical Aspects of Computer Science (STACS)*, volume 2607 of *LNCS*, pages 675–686. Springer, 2003.
36. Michael Backes, Birgit Pfitzmann, Michael Steiner, and Michael Waidner. Polynomial fairness and liveness. In *Proceedings of 15th IEEE Computer Security Foundations Workshop (CSFW)*, pages 160–174, 2002.
37. Michael Backes and Birgit Pfitzmann. Intransitive non-interference for cryptographic purposes. In *Proc. 24th IEEE Symp. on Security & Privacy*, pages 140–152, 2003.
38. Michael Backes. Quantifying probabilistic information flow in computational reactive systems. In *Proceedings of 10th European Symposium on Research in Computer Security (ESORICS)*, volume 3679 of *Lecture Notes in Computer Science*, pages 336–354. Springer, 2005.
39. Daniele Micciancio and Bogdan Warinschi. Soundness of formal encryption in the presence of active adversaries. In *Proc. 1st TCC*, volume 2951 of *LNCS*, pages 133–151. Springer, 2004.
40. Peeter Laud. Symmetric encryption in automatic analyses for confidentiality against active adversaries. In *Proc. 25th IEEE Symp. on Security & Privacy*, pages 71–85, 2004.
41. Ran Canetti and Jonathan Herzog. Universally composable symbolic analysis of mutual authentication and key exchange protocols. In *Proc. 3rd TCC*, pages 380–403. Springer, 2006.
42. Vèronique Cortier and Bogdan Warinschi. Computationally sound, automated proofs for security protocols. In *Proc. 14th ESOP*, pages 157–171, 2005.
43. M. Baudet, V. Cortier, and S. Kremer. Computationally sound implementations of equational theories against passive adversaries. In *Proc. 32nd ICALP*, volume 3580 of *LNCS*, pages 652–663. Springer, 2005.
44. Michael Backes, Christian Jacobi, and Birgit Pfitzmann. Deriving cryptographically sound implementations using composition and formally verified bisimulation. In *Proceedings of 11th International Symposium on Formal Methods Europe (FME)*, volume 2391 of *Lecture Notes in Computer Science*, pages 310–329. Springer, 2002.
45. Michael Backes, Birgit Pfitzmann, and Michael Waidner. Low-level ideal signatures and general integrity idealization. In *Proceedings of 7th Information Security Conference (ISC)*, volume 3225 of *Lecture Notes in Computer Science*, pages 39–51. Springer, 2004.

46. Michael Backes and Birgit Pfitzmann. Limits of the cryptographic realization of Dolev-Yao-style XOR. In *Proceedings of 10th European Symposium on Research in Computer Security (ESORICS)*, volume 3679 of *Lecture Notes in Computer Science*, pages 178–196. Springer, 2005.
47. Michael Backes, Markus Duermuth, Dennis Hofheinz, and Ralf Kuesters. Conditional reactive simulatability. In *Proceedings of 11th European Symposium on Research in Computer Security (ESORICS)*, volume 4189 of *Lecture Notes in Computer Science*, pages 424–443. Springer, 2006. Preprint on IACR ePrint 2006/132.
48. Michael Backes, Birgit Pfitzmann, and Michael Waidner. Limits of the reactive simulatability/UC of Dolev-Yao models with hashes. In *Proceedings of 11th European Symposium on Research in Computer Security(ESORICS)*, volume 4189 of *Lecture Notes in Computer Science*, pages 404–423. Springer, 2006.
49. Michael Backes, Markus Dürmuth, and Ralf Küsters. On simulatability soundness and mapping soundness of symbolic cryptography. In *Proceedings of 27th International Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, 2007.
50. Michael Backes and Dominique Unruh. Computational soundness of symbolic zero-knowledge proofs against active attackers. In *21st IEEE Computer Security Foundations Symposium, CSF 2008*, pages 255–269, 2008. Preprint on IACR ePrint 2008/152.
51. Michael Backes and Birgit Pfitzmann. Relating symbolic and cryptographic secrecy. *IEEE Transactions on Dependable and Secure Computing*, 2(2):109–123, 2005.
52. J. Mitchell, M. Mitchell, and A. Scedrov. A linguistic characterization of bounded oracle computation and probabilistic polynomial time. In *Proc. 39th FOCS*, pages 725–733, 1998.
53. P. Lincoln, J. Mitchell, M. Mitchell, and A. Scedrov. A probabilistic poly-time framework for protocol analysis. In *Proc. 5th ACM CCS*, pages 112–121, 1998.
54. Russell Impagliazzo and Bruce M. Kapron. Logics for reasoning about cryptographic constructions. In *Proc. 44th IEEE FOCS*, pages 372–381, 2003.
55. Anupam Datta, Ante Derek, John Mitchell, Vitalij Shmatikov, and Matthieu Turuani. Probabilistic polynomial-time semantics for a protocol security logic. In *Proc. 32nd ICALP*, volume 3580 of *LNCS*, pages 16–29. Springer, 2005.
56. Bruno Blanchet. A computationally sound mechanized prover for security protocols. In *Proc. 27th IEEE Symp. on Security & Privacy*, pages 140–154, 2006.
57. Martín Abadi. Secrecy by Typing in Security Protocols. *Journal of the ACM*, 46(5):749–786, September 1999.
58. Martín Abadi and Bruno Blanchet. Analyzing Security Protocols with Secrecy Types and Logic Programs. *Journal of the ACM*, 52(1):102–146, January 2005.
59. Andrew D. Gordon and Alan Jeffrey. Authenticity by Typing for Security Protocols. *Journal of Computer Security*, 11(4):451–520, 2003.
60. Andrew D. Gordon and Alan Jeffrey. Typing correspondence assertions for communication protocols. *Theoretical Computer Science*, 300(1-3):379–409, 7 May 2003.
61. Andrew D. Gordon and Alan Jeffrey. Types and effects for asymmetric cryptographic protocols. *Journal of Computer Security*, 12(3-4):435–483, 2004.
62. Peeter Laud. Handling Encryption in Analyses for Secure Information Flow. In *Proc. ESOP 2003*, volume 2618 of *LNCS*, pages 159–173. Springer, 2003.
63. Peeter Laud and Varmo Vene. A Type System for Computationally Secure Information Flow. In *Proc. 15th FCT*, volume 3623 of *LNCS*, pages 365–377. Springer, 2005.
64. Chiara Bodei, Mikael Buchholtz, Pierpaolo Degano, Flemming Nielson, and Hanne Riis Nielson. Static Validation of Security Protocols. *Journal of Computer Security*, 13(3):347–390, 2005.
65. Christian Fecht and Helmut Seidl. An Even Faster Solver for General Systems of Equations. In *Proc. 3rd SAS*, volume 1145 of *LNCS*, pages 189–204. Springer, 1996.

66. Roger M. Needham and Michael D. Schroeder. Using Encryption for Authentication in Large Networks of Computers. *Communications of the ACM*, 21(12):993–999, December 1978.
67. D. Otway and O. Rees. Efficient and timely mutual authentication. *Operating Systems Review*, 21(1):8–10, 1987.
68. Michael Burrows, Martín Abadi, and Roger M. Needham. A Logic of Authentication. *ACM Transactions on Computer Systems*, 8(1):18–36, February 1990.
69. Paul F. Syverson. A Taxonomy of Replay Attacks. In *Proceedings of the 7th IEEE CSFW*, pages 187–191, 1994.
70. Michael Backes and Peeter Laud. Computationally Sound Secrecy Proofs by Mechanized Flow Analysis. Cryptology ePrint Archive: Report 2006/266, 10 August 2006.
71. Vesal Vojdani. Mitmelõimeliste C-programmide kraasimine analüsaatoriga Goblin (Linting multi-threaded C programs with the Goblin). Master's thesis, Tartu University, 2006.