Reasoning about Belief in Cryptographic Protocols*

Li Gong, Roger Needham, and Raphael Yahalom†

University of Cambridge Computer Laboratory
Cambridge CB2 3QG, England

February 11, 1990

Abstract

Abstract. Analysis methods for cryptographic protocols have often focused on information leakage rather than on seeing whether a protocol meets its goals. Many protocols, however, fall far short of meeting their goals, sometimes for quite subtle reasons. We introduce a mechanism for reasoning about belief as a systematic way to understand the working of cryptographic protocols. Our mechanism captures more features of such protocols than that given in a recent work [1], to which our proposals are a substantial extension.

1 Introduction

Solutions to computer security problems over the last few years have brought forth the need for rigorous analysis methods. Formal tools must be provided to determine whether a solution indeed solves a problem, as well as to enable comparisons between proposed solutions. In this paper we propose a method for reasoning about cryptographic protocols in a distributed environment.

The work described was inspired by the recent development of a modal logic to reason about authentication protocols [1], which we refer to as the BAN logic. Indeed our work can be seen as a new approach within the framework proposed there. Like BAN, we aim to analyze a protocol step by step, make explicit any assumptions required, and draw conclusions about the final position it attains.

Our new approach seems to offer important advantages over the BAN approach. It does not require several universal assumptions which the BAN work does. For example, it does not assume that redundancy is always present in encrypted messages - incorporating instead a new notion of *recognizability* which captures a recipient's expectation of the contents of messages he receives. Also, it does not assume that a principal can always determine whether a message was not once originated by himself.

We distinguish between what one possesses and what one believes in. This allows us to treat separately the content of a message and the information *implied* by such a message. It also makes it possible to separate reasoning about the physical world from the reasoning about other principals' beliefs, so that we can consider different levels of trust in the reasoning.

The set of notions is expanded so that additional properties of messages can be incorporated in the reasoning process. Some of the existing notions are modified and made to correspond more naturally to execution states and are thus more intuitive. For example, encrypted and plaintext formulae are treated similarly, thus permitting reasoning about multiply encrypted messages. Also, plaintext messages, which are not considered in the BAN approach to be useful, can in some cases be used to derive further conclusions.

The new approach allows us to analyze a much wider range of protocols. A summary of the significant differences between our work and the BAN work appears at the end of the paper.

The rest of this paper is organized as follows. In the next section we outline the model of computation. In section 3 we introduce the notions and their corresponding notation. In section 4 we present selected logical postulates underlying our reasoning process. The complete set of postulates and their descriptions are included in appendix A. In section 5 our reasoning process is described. In section 6 we introduce reasoning about other principals' beliefs. In section 7 the Needham-Schroeder protocol is used as an example that highlights some of the characteristics of our approach. Finally, in section 8 we discuss our conclusions.

2 The Model of Computation

Our model of computation is similar to that used in the BAN work and possesses some characteristics of models used in other knowledge-theoretic work (e.g. [2]). We

*Published in Proceedings of the IEEE 1990 Symposium on Security and Privacy, Oakland, California, May, 1990, pp.234-248.

†Raphael Yahalom is supported by IBM's UK Academic Systems Marketing and by a British government's Overseas Research Students award.

outline the significant aspects of the model. A more formal presentation is given in appendix C.

A distributed environment consists of principals, essentially state-machines, which are connected by communication links. Messages on these links constitute the only means of communication between principals. Any principal can place a message on any link. He can also see and alter any message being passed along any link.

A protocol is a distributed algorithm. A protocol determines what messages should be sent by the participating principals as a function of their internal states. A run is a particular execution of the protocol. We refer to a protocol run as a session.

Each principal in each session maintains two sets: a *belief* set which includes all the current beliefs of the principal, and a *possession* set which includes all the formulae available to the principal. In particular, the latter includes everything the principal received, and everything the principal has generated himself, e.g. random numbers, in the current session.

Principals start a session with certain initial beliefs and initial possessions. From that point a principal can obtain new beliefs, and thus expand his belief set, as a result of receiving new messages. Inference rules, described below, enable the derivation of new beliefs from current beliefs and incoming messages. Similarly, a principal can increase his possessions.

Beliefs and possessions are monotonic within a given session. That is to say, if a belief or a possession is a member of its respective set at any phase of some session, then it is a member of that set at any subsequent phase of that session. However, no such claim is made across sessions. In particular, a belief or a possession of a principal in a session is not necessarily a member of the corresponding set in future, or past, sessions in which the principal participates.

The only universal assumption we require is that principals do not reveal their secrets. This is, in fact, only for convenience. We could let principals choose, as part of their initial beliefs, whether to trust each other in that respect. Without such basic trust, however, a principal would be rather limited in his ability to attain new beliefs.

3 Notions and Notation

In this section we introduce the basic notions underlying our reasoning process and their corresponding notation. The description below is informal and aims at providing intuitive understanding. The postulates presented in the next section, and the semantics in appendix C, provide a more precise definition.

3.1 Formulae

A formula is a name used to refer to a bit string, which would have a particular value in a run. This is rather like the name of a variable. Let X and Y range over formulae. Two kinds of special formulae, shared secrets and encryption keys, are denoted as S and K respectively. The following are also formulae:

(X, Y): conjunction of two formulae. We treat conjunctions as sets with properties such as associativity and commutativity.

$\{X\}_K$ and $\{\mathrm{X}\}_K^{-1}$: conventional encryption and decryption (e.g. DES). It is assumed that the cryptosystems used are resistant to ciphertext-only and known-plaintext attacks. In addition, every bit in a ciphertext depends on all bits of the plaintext and the key in such a way that any change to the plaintext causes a random change in the ciphertext and vice versa. They satisfy $\{\{X\}_K\}_K^{-1} = X$, but not necessarily $\{\{X\}_K^{-1}\}_K = X$.

$\{X\}_{+K}$ and $\{X\}_{-K}$: public-key encryption and decryption. In addition to the requirements stated for conventional cryptosystems, they satisfy $\{\{X\}_{+K}\}_{-K} = X$. Some public-key schemes (e.g. RSA [8]) also satisfy that $\{\{X\}_{-K}\}_{+K} = X$.

$H(X)$: a one-way function of X. It is required that given X it is computationally feasible to compute $H(X)$; given $H(X)$ it is infeasible to compute X; it is infeasible to compute X and X' such that $X \neq X'$ but $H(X) = H(X')$ [5].

$F(X_1, \ldots, X_n)$: F is a many-to-one computationally feasible
function such that for any X_i, $1 \leq i \leq n$, and constants $C_1, \ldots, C_{n-1}$, $F(C_1, \ldots, C_{i-1}, X_i, C_i, \ldots, C_{n-1})$ is a one to one computationally feasible function, and its inverse is also computationally feasible. For example, exclusive-or is such a function. $\bar{F}(X)$ denotes a computationally feasible one-to-one function whose inverse is also computationally feasible.

3.2 Statements

A basic statement reflects some property of a formula. Let P and Q range over principals. The following are basic statements.

$P \lhd X$: P *is told* formula X. P receives X, possibly after performing some computation such as decryption. That is, a formula being told can be the message itself, as well as any computable content of that message.

$P \ni X$: P *possesses*, or is capable of possessing, formula X. At a particular stage of a run, this includes all the formulae that P has been told, all the formulae he started the session with, and all the ones he has generated in that run. In addition P possesses, or is capable of possessing, everything that is computable from the formulae he already possesses.

$P \mid\sim X$: P *once conveyed* formula X. X can be a message itself or some content computable from such a message, i.e. a formula can be conveyed implicitly.

$P \mid\equiv \sharp(X)$: P *believes*, or is entitled to believe, that formula X is *fresh*. That is, X has not been used for the same purpose at any time before the current run of the protocol. For example, a counter or a random number generator (of sufficient quality) can serve to produce formulae that a principal believes to be fresh (called *nonces* [6]).

$P \mid\equiv \phi(X)$: P *believes*, or is entitled to believe, that formula X is *recognizable*. That is, P would recognize

X if P has certain expectations about the contents of X before actually receiving X. P may recognize a particular value (e.g. his own identifier), a particular structure (e.g. the format of a timestamp), or a particular form of redundancy.

$P \mid\equiv P \overset{S}{\leftrightarrow} Q$: P *believes*, or is entitled to believe, that S is a suitable *secret* for P and Q. They may properly use it to mutually prove identity. They may also use it as, or derive from it, a key to communicate. S will never be discovered by any principal except P, Q, or a principal trusted by either P or Q. However, in this case, the trusted principal should never use S as proof of identity or as a key to communicate. This notation is symmetrical: $Q \overset{S}{\leftrightarrow} P$ and $P \overset{S}{\leftrightarrow} Q$ can be used interchangeably.

$P \mid\equiv \overset{+K}{\mapsto} Q$: P *believes*, or is entitled to believe, that $+K$ is a suitable *public key* for Q. The matching secret key $-K$ will never be discovered by any principal except Q or a principal trusted by Q. In this case, however, the trusted principal should not use it to prove identity or to communicate.

Let C range over statements. The following are also statements:

C_1, C_2: conjunction. We treat conjunctions as sets with properties such as associativity and commutativity.

$P \mid\equiv C$: P believes, or P would be entitled to believe, that statement C holds.

3.3 "Not-Originated-Here" Formulae

A formula can be regarded as a *not-originated-here* formula, denoted by prefixing a star to the formula, e.g. $*X$. Statement $P \lhd *X$ indicates that P is told a formula which he did not convey previously in the current run.

Suppose a shared secret key K is used between P and Q. If P constructs and conveys a formula $\{X\}_K$ and P is later told the exact packet, possibly encrypted under other keys, then P should not reason that Q once conveyed X. The reason is that P could have sent it himself. A not-originated-here formula implies that it was not first conveyed by the recipient in this session (but could have been conveyed by him in previous sessions). A principal can believe he has never conveyed a formula that he receives if it is a not-originated-here formula and he also believes the formula to be fresh.

We use a pattern scanner to screen a normal protocol description and insert those stars at the correct places. The algorithm is discussed in section 5.

In appendix B we discuss reasoning in environments where principals may believe that they can identify messages that were not originated by themselves in any session.

4 Logical Postulates

In this section we introduce the logical postulates underlying the reasoning process. There are five categories of postulates. We describe each category and present representative postulates. A complete list of all the logical postulates and their description is included in appendix A.

4.1 Being-Told Rules

The first set of rules deals with formulae a principal receives. We regard every formula a principal receives, as well as certain manipulations of that formula (e.g. decryption with certain keys), as *being told* to that principal. The following are examples of *being-told* rules:

T2 $$\frac{P \lhd (X, Y)}{P \lhd X}$$

Being told a formula implies being told each of its concatenated components.

T3 $$\frac{P \lhd \{X\}_K,\ P \ni \mathrm{K}}{P \lhd X}$$

If a principal is told a formula encrypted with a key he possesses then he is also considered to have been told the decrypted contents of that formula.

4.2 Possession Rules

The next set of rules specify the formulae a principal is capable of possessing by manipulating formulae he already possesses. Examples of such rules are:

P1 $$\frac{P \lhd X}{P \ni X}$$

A principal is capable of possessing anything he is told.

P2 $$\frac{P \ni X,\ P \ni Y}{P \ni (X, Y),\ P \ni F(X, Y)}$$

If a principal possesses two formulae then he is capable of possessing the concatenation of the two formulae and a function F of them.

4.3 Freshness Rules

Freshness rules specify the formulae a principal can believe to be fresh, given his beliefs about the freshness of other formulae. Recall that a principal's belief in the freshness of a formula represents his belief that the formula has never had the same value in any previous run of the protocol.

F1 $$\frac{P \mid\equiv \sharp(X)}{P \mid\equiv \sharp(X, Y),\ P \mid\equiv \sharp(F(X))}$$

If P believes a formula X is fresh, then he is entitled to believe that any formula of which X is a component is fresh, and that a computationally feasible one-to-one function F of X is fresh. For convenience, we use $P \mid\equiv \sharp(X, Y)$ to denote $P \mid\equiv \sharp(X)$ or $P \mid\equiv \sharp(Y)$.

F2 $$\frac{P \mid\equiv \sharp(X),\ P \ni K}{P \mid\equiv \sharp(\{X\}_K),\ P \mid\equiv \sharp(\{X\}_K^{-1})}$$

If P believes a formula X is fresh and possesses a key, then P is entitled to believe that the encryption, as well as the decryption, of X with that key is fresh.

4.4 Recognizability Rules

Recognizability rules specify the formulae a principal can believe to be recognizable, given his beliefs about the recognizability of other formulae. Examples of recognizability rules are:

R1 $$\frac{P \mid\equiv \phi(X)}{P \mid\equiv \phi(X,Y),\ P \mid\equiv \phi(F(X))}$$

If a principal P believes a formula X is recognizable, then he is entitled to believe that any formula of which X is a component is recognizable, and that a computationally feasible function F of X is recognizable.

R2 $$\frac{P \mid\equiv \phi(X),\ P \ni K}{P \mid\equiv \phi(\{X\}_K),\ P \mid\equiv \phi(\{X\}_K^{-1})}$$

If a principal P believes a formula X is recognizable, and P possesses a key K, then P is entitled to believe that the encryption and the decryption of X with K are recognizable.

4.5 Message Interpretation Rules

Finally, we have rules which enable principals to advance their beliefs about other principals by examining messages they receive.

Postulate I1 below (as well as I2 and I3 in appendix A) includes a freshness requirement that may seem rather surprising. It is a consequence of the fact that when a key is shared between two (or more) principals, each could have used the secret to construct a formula. A principal should only be convinced that a message based on a shared secret is not a replay of one of his own messages if the message has a not-originated-here star associated with it, and he believes the message is fresh.

I1 $$\frac{\begin{array}{c}P \lhd *\{X\}_K,\ P \ni K,\ P \mid\equiv P \overset{K}{\leftrightarrow} Q,\\ P \mid\equiv \phi(X),\ P \mid\equiv \sharp(X,K)\end{array}}{P \mid\equiv Q \mid\sim X,\ P \mid\equiv Q \mid\sim \{X\}_K,\ P \mid\equiv Q \ni K}$$

If for a principal P, all of the following conditions hold: (1) P receives a formula consisting of X encrypted with K and marked with a not-originated-here mark; (2) P possesses key K; (3) P believes K is a suitable secret for himself and Q; (4) P believes formula X is recognizable; (5) P believes that K or X are fresh. Then P is entitled to believe the following: Q once conveyed X; Q once conveyed the formula X encrypted with K; and Q possesses K.

I6 is an example of rules that permit reasoning about the state of a sender:

I6 $$\frac{P \mid\equiv Q \mid\sim X,\ P \mid\equiv \sharp(X)}{P \mid\equiv Q \ni X}$$

If P believes that Q once conveyed formula X and P believes that X is fresh, then P is entitled to believe that Q possesses X.

4.6 Rationality Rule

We supplement the postulates with a rule that we call the rationality rule. Informally, it states that our set of postulates can be expanded to permit reasoning about a principal's beliefs regarding the state of other principals. More precisely:

if $\frac{C1}{C2}$ is a postulate, then for any principal P, so is $\frac{P \mid\equiv C1}{P \mid\equiv C2}$.

For example, from postulate P2 we can obtain the following postulate:

$$\frac{Q \mid\equiv P \ni X,\ Q \mid\equiv P \ni Y}{Q \mid\equiv P \ni F(X,Y)}$$

If Q believes that P is capable of possessing two formulae then Q believes that P is capable of possessing the concatenation of the two formulae and a function F of them.

The rationality rule represents our view that principals are capable of deriving rational conclusions about the state of other principals. Issues associated with reasoning about other principals' beliefs are discussed in section 6.

5 Protocol Analysis

5.1 A Protocol Parser

Protocols are typically described by listing messages sent between the principals, and by symbolically showing the source, the destination, and the contents of each message. This form is intuitively easy to understand. We require only simple transformations to attain a form suitable for direct manipulation in our logic.

The main point is related to the fact that a typical protocol description does not make a distinction between X and $*X$. $P \lhd *X$ denotes the fact that a P is not the first one to convey X in the current run of the protocol, a fact that is only implicitly included in the description itself. We design a parser that explicitly inserts the stars to a protocol description, thus avoiding a much more complex form of logic that would have otherwise been required. We sketch out a general description of the parser algorithm.

For each line from the description $P \rightarrow Q : X$, if $Q = P$ an error is reported; otherwise, the parser produces two lines, $P \mid\sim X$ followed by Q $\lhd X$. For each principal P, the parser examines all lines of the form $P \mid\sim X$ or $P \lhd X$. It then scans from the beginning. For each

pattern of a complete formula Y in a line $P \lhd X$, if Y does not first appear in a line $P \mid\sim X$ in the protocol, the parser inserts a star before Y. The parser would also mark $(*X, *Y)$ instead of $*(X, Y)$. After these are completed, the parser drops all lines of the form $P \mid\sim X$.

In addition, for private-key encryption, the parser replaces patterns of the form $\{\{X\}_K\}_K^{-1}$ with X. Similar replacements are performed for public-key encryption.

5.2 Annotated Assertions

The protocol analysis consists of annotating the protocols with statements and manipulating these statements with the postulates. A protocol is a sequence of *told*-statements $C_1, \cdots, C_n$, each of the form $P \lhd X$. An annotation for a protocol consists of a sequence of assertions, conjunctions of statements, inserted before the first told-statement and after each told-statement. The first assertion contains the assumptions and the last contains the conclusions. They can be understood as formulae in Hoare logic [3]. As in BAN [1], if the assumptions hold, each assertion should hold after the execution of its respective protocol prefix. The assertions are derived by the syntactic application of the above postulates to statements. Often, the goal of an analysis is to derive the final positions of each of the principals at the end of a given protocol. That final position is represented by the corresponding assertions.

5.3 Example - A Voting Protocol

We present a simple voting protocol, outline its characteristics, and demonstrate the reasoning process as described thus far.

The environment consists of a coordinator Q and n participants P_i. Communication is performed by message passing. Each participant P_i shares a secret with Q, denoted as S_i. The goal of the protocol is to elect a principal for some position. Each participant has one vote V_i. The coordinator determines the winner according to certain rules.

The protocol requirements specify that the coordinator should be able to identify votes for the current election and that each principal who possesses the result R at the end of the protocol should be convinced that it was generated by the coordinator during the current run.

The protocol consists of three messages between the coordinator and each participant. These correspond to request-vote, voting, and result-announcement phases. N_q, and N_i are nonces generated by Q and P_i respectively. A secret S used for identification purposes is denoted $< S >$ (see appendix A.5).

1. $Q \rightarrow P_i$: N_q
2. $P_i \rightarrow Q$: $P_i, N_i, V_i, H(N_q, < S_i >, V_i)$
3. $Q \rightarrow P_i$: $R, H(N_i, < S_i >, R)$

The parser algorithm would produce the following description of the protocol.

1. $P_i \lhd$: $*N_q$
2. $Q \lhd$: $*P_i, *N_i, *V_i, *H(N_q, * < S_i >, V_i)$
3. $P_i \lhd$: $*R, *H(N_i, < S_i >, R)$

5.4 Protocol Analysis

We assume that the following holds at the beginning of every run of the protocol:

$$P_i \ni S_i; \quad P_i \ni N_i; \quad P_i \mid\equiv Q \overset{S_i}{\leftrightarrow} P_i; \quad P_i \mid\equiv \sharp(N_i)$$

$$Q \ni S_i; \quad Q \ni N_q; \quad Q \mid\equiv Q \overset{S_i}{\leftrightarrow} P_i; \quad Q \mid\equiv \sharp(N_q)$$

That is, each one of the participants possesses a secret and believes that it is for himself and the coordinator. Each also possesses a nonce and believes in its freshness. Similarly the coordinator possesses the secrets and believes each for him and a principal. He also possesses a nonce and believes it is fresh.

For any run of the protocol:

Message 1: Applying T1 and P1 we obtain $P_i \ni N_q$. That is P_i possesses N_q.

Message 2: Applying T1, T2, and P1 we obtain $Q \ni (V_i, N_i)$. Q possesses V_i and N_i.

Applying F1 we obtain $Q \mid\equiv \sharp(V_i, N_q, S_i)$. Q believes that (V_i, N_q, S_i) is fresh. That is, it is generated during the current run of the protocol and so cannot be a replay of a message from previous runs.

Applying I3 and I7 we obtain $Q \mid\equiv P_i \mid\sim (V_i, N_q)$. Applying F1 we obtain $Q \mid\equiv \sharp(V_i, N_q)$. Q believes that P_i conveyed (V_i, N_q) in the current run.

Similarly, in message 3, P_i believes that Q conveyed R in the current run.

The final position that the coordinator and each of the participant attain correspond to the specification goals outlined earlier. In particular, the coordinator is in a position to recognize fresh and valid vote messages from all participants and to properly determine the winner. The participants possess a result which was generated by the coordinator in the current session.

The reasoning process based on the postulates above can be considered a *physical level* reasoning process. Each principal can only advance his beliefs and increase his possessions based on the physical content of the messages he receives. Consequently a principal can attain beliefs about who sent what and when, but not about the sender's beliefs at the time the message was sent. In the voting protocol the participants are convinced that the result was originated by the coordinator, but there is no way of deriving conclusions such as that the coordinator himself believes in the validity of the result. Thus far, our formal protocol description did not include such notions. They are discussed in the next section.

6 Beliefs about Others' Beliefs

The reasoning process described thus far permits conclusions of the following kind: $P \mid\equiv Q \ni X$, $P \mid\equiv Q \mid\sim$

X, $P \mid\equiv \sharp(X)$, and so on. These are principal P's beliefs about the physical world. However, it does not allow for conclusions such as $P \mid\equiv Q \mid\equiv \sharp(X)$, i.e. P cannot draw any conclusion about beliefs held by other principals.

We choose to separate these two types of beliefs for several reasons. First, by examining the contents of messages one can, quite directly, derive conclusions about the possessor of a formula or the conveyor of a message, that is, reason about the physical world. At that level there is no need for reasoning about other principals' beliefs.

Secondly, since each principal expresses his beliefs by sending messages, we choose to interpret the beliefs on the sender's part as the preconditions for a message to be sent at all. That is, the recipient should interpret it as such when receiving the message. A message of the same form may carry different, context dependent, meanings in different protocols. Having believed a message to be genuine, a recipient can choose to believe the sender's beliefs, if he trusts the sender's honesty and competence in following the protocol specification. In other words, since we do not require the universal assumption that all principals are honest and competent, we should reason about beliefs held by others based on trust of different levels.

Thirdly, typical protocol specifications often include verbal description to the effect that a principal should proceed only if certain conditions hold or only if he holds certain beliefs. This can be regarded as a precondition. Thus we feel that our approach is natural.

The precondition of a formula X being conveyed, represented by statement C, is described in the protocol as $X \rightsquigarrow C$. We call C a *message extension*. Since a message can contain a number of formulae that are destined for different principals and possibly with different meanings, we use $P \triangleleft (X_1 \rightsquigarrow C_1), (X_2 \rightsquigarrow C_2)$ to clearly express the scope.

Now recall that the parser may find a pattern that first appears in a line of the form $P \mid\sim X$. This may indicate that a new formula is conveyed by P under a certain precondition. We shall take, for example, the verbal explanation often found in protocol descriptions, translate this into the language of our logic, and insert it after the formula. The same condition should be replicated whenever the same pattern reappears in the following messages. As before, only lines of the form $P \triangleleft (X \rightsquigarrow C)$ are kept. For simplicity, a formula without a star prefix has no conditions attached, since no principal should derive new conclusions from such a formula. The extension of a message is considered part of that message, and so the conclusion that a principal once conveyed a formula can be augmented to the conclusion that he conveyed the formula and its extension - if such an extension is associated with the formula.

6.1 Trust and Jurisdiction

We use the notion of *jurisdiction* to represent trust and delegation.

$P \mid\equiv Q \mid\Longrightarrow C$: P believes that Q has jurisdiction over statement C. That is, P believes that principal Q is an authority on C and should be trusted on this matter.

The following postulate describes jurisdiction more formally:

$$\textbf{J1} \qquad \frac{P \mid\equiv Q \mid\Longrightarrow C,\ P \mid\equiv Q \mid\equiv C}{P \mid\equiv C}$$

J1 states that if P believes that Q has jurisdiction over some statement C and that Q believes in C, then P ought to believe in C as well. The following construct is a special case of jurisdiction.

$P \mid\equiv Q \mid\Longrightarrow Q \mid\equiv * :$ P believes that Q has jurisdiction over all his beliefs. The principal Q is considered by P to be completely honest and competent.

J2 states that if P believes that Q is honest and competent, and P receives a message $X \rightsquigarrow C$, which is originated from Q, then P ought to believe that Q really believes C. J3 is a special case.

$$\textbf{J2} \qquad \frac{P \mid\equiv Q \mid\Longrightarrow Q \mid\equiv *,\ P \mid\equiv Q \mid\sim (X \rightsquigarrow C),\ P \mid\equiv \sharp(X)}{P \mid\equiv Q \mid\equiv C}$$

$$\textbf{J3} \qquad \frac{P \mid\equiv Q \mid\Longrightarrow Q \mid\equiv *,\ P \mid\equiv Q \mid\equiv Q \mid\equiv C}{P \mid\equiv Q \mid\equiv C}$$

Other levels of trust could be defined and used for reasoning. For example, one could introduce a notion representing the fact that a principal is following the protocol. In other words, the relationship between the contents of the messages he sends and their *meaning* is maintained. Consequently, a similar postulate to J2, based on that notion, would conclude $P \mid\equiv Q \mid\sim Q \mid\equiv C$ rather than $P \mid\equiv Q \mid\equiv C$. Clearly, this represents a lower level of trust and so enables weaker conclusions to be derived. Moreover, any of the above levels of trust could be applied to any aspect of a principal's behavior. This reflects the fact that principals are often trusted differently with respect to the different tasks they perform.

Recall that the rationality rule introduced in section 4 enables principals to further increase their beliefs concerning other principals' beliefs.

6.2 Consistency of Protocol Description

Our approach to reasoning about principals' states enables us to detect invalid or inconsistent protocol descriptions. In particular, a principal should include in a message only formulae he possesses at the time the message is constructed. Similarly, a principal should imply only his beliefs which, according to our reasoning, he holds at the time the message is sent. Note that present discussion refers to the protocol description - during protocol execution principals may lie.

The following two checks are thus performed: *possession consistency*, i.e. a principal should only be able to include in any message he sends, a formula he possesses; *belief consistency*, i.e. a message extension should include only beliefs held by the sender at the time the message is sent.

7 The Needham-Schroeder Protocol

In this section we apply the reasoning process to the Needham-Schroeder authentication protocol [6]. We choose this protocol as an example because it influenced the design of a significant number of existing systems and published protocols, and because it serves as one of the examples in the BAN logic paper [1]. Our analysis below can thus provide insight to some of the differences between the approaches, and to the advantages that ours offers. In the rest of this section we describe the protocol, analyze it, and finally investigate how modifications to the protocol affect the final position of the participants.

7.1 Protocol Description

The general goal of the protocol is for two principals P and Q to be provided with a shared secret [6, 1]. That secret can consequently be used as a session key. There exists a trusted *authentication server* S, which shares common secrets with all potential participants and can generate good quality sessions keys.

Different authentication protocols may differ in the final positions which the principals attain. Different positions seem to suffice for different environments. It is usually assumed that the minimum requirement for an end of a run is that each principal would possess a session key, and be convinced of the validity and quality of that key. Often, in addition, each principal is required to believe something about the state of the other principal. That can vary between believing that the other principal is operational, that the other principal possesses the key, or that the other principal believes in the validity of the possessed key.

The Needham-Schroeder protocol consists of the following five messages:

1. $P \rightarrow S$: P, Q, N_p
2. $S \rightarrow P$: $\{N_p, Q, K, \{K, P\}_{K_{qs}}\}_{K_{ps}}$
3. $P \rightarrow Q$: $\{K, P\}_{K_{qs}}$
4. $Q \rightarrow P$: $\{N_q\}_K$
5. $P \rightarrow Q$: $\{N_q - 1\}_K$

N_p and N_q are nonces for P and Q respectively, K_{ps} and K_{qs} are the secrets between P and S, Q and S respectively. K is the session key for P and Q generated by S. To the above description, we add message extensions which reflects the verbal explanation of the protocol execution. The validity of these extensions will be discussed during the analysis of the protocol. We use F to denote the decrement computation. The parser would produce the following output.

1. $S\triangleleft$: $*P, *Q, *N_p$
2. $P\triangleleft$: $*\{N_p, Q, *K, *\{K, P\}_{K_{qs}} \rightsquigarrow S \mid\equiv P \overset{K}{\leftrightarrow} Q\}_{K_{ps}}$ $\rightsquigarrow S \mid\equiv P \overset{K}{\leftrightarrow} Q$
3. $Q\triangleleft$: $*\{*K, *P\}_{K_{qs}} \rightsquigarrow S \mid\equiv P \overset{K}{\leftrightarrow} Q$
4. $P\triangleleft$: $*\{*N_q\}_K$
5. $Q\triangleleft$: $*\{*F(N_q)\}_K \rightsquigarrow P \mid\equiv P \overset{K}{\leftrightarrow} Q$

7.2 Protocol Analysis

We begin by listing the initial assumptions:

$P \ni K_{ps}; \quad P \ni N_p$

$P \mid\equiv P \overset{K_{ps}}{\leftrightarrow} S; \quad P \mid\equiv \sharp(N_p); \quad P \mid\equiv \phi(Q)$

$Q \ni K_{qs}; \quad Q \ni N_q$

$Q \mid\equiv Q \overset{K_{qs}}{\leftrightarrow} S; \quad Q \mid\equiv \sharp(N_q); \quad Q \mid\equiv \phi(N_q)$

That is, each principal possesses a secret and believes it is a secret between himself and the authentication server. He also possesses a nonce which he believes to be fresh. In addition, P believes that the identifier Q is recognizable and Q believes that N_q is recognizable.

$P \mid\equiv S \Longrightarrow (P \overset{K}{\leftrightarrow} Q); \quad P \mid\equiv S \Longrightarrow S \mid\equiv *;$
$P \mid\equiv Q \Longrightarrow Q \mid\equiv *$

$Q \mid\equiv S \Longrightarrow (P \overset{K}{\leftrightarrow} Q); \quad Q \mid\equiv S \Longrightarrow S \mid\equiv *;$
$Q \mid\equiv P \Longrightarrow P \mid\equiv *$

P and Q believe in the jurisdiction of S over quality secrets to be shared between them. P and Q also believe that S is honest and competent. Moreover, P believes Q is competent and honest, and vice versa. These two assumptions are not essential. The protocol could attain a useful final position even if the principals do not completely trust each other to tell the truth. However, as we shall see, the added trust represented by these last two assumptions enables both P and Q to attain stronger final positions.

$S \ni K_{ps}; \quad S \ni K_{qs}; \quad S \ni K$

$S \mid\equiv P \overset{K_{ps}}{\leftrightarrow} S; \quad S \mid\equiv Q \overset{K_{qs}}{\leftrightarrow} S; \quad S \mid\equiv P \overset{K}{\leftrightarrow} Q$

S believes that he possesses valid keys with P and Q. He also believes that K is a suitable secret for P and Q.

For any run of the protocol:

Message 1: applying T1 and P1 we obtain $S \ni (P, Q, N_p)$. That is S possesses P, Q, and N_p.

Message 2: first we note that the extension to the message, $S \mid\equiv P \overset{K}{\leftrightarrow} Q$, is valid because it holds when the message is sent as is evident from the initial assumptions. Also S is also sure that the recipient, P, cannot mistake K as a key for himself and a principal other than Q, since the name Q is included in the message.

Applying T1, T3, and P1 we get $P \ni (N_p, Q, K, \{K, P\}_{K_{qs}})$. P possesses the message contents.

Applying T2 we obtain $P \ni K$. P possesses the new key K.

Applying F1 we obtain $P \mid\equiv \sharp(N_p, Q, K, \{K, P\}_{K_{qs}})$. P believes the message is not a replay.

Applying R1 we obtain $P \mid\equiv \phi(N_p, Q, K, \{K, P\}_{K_{qs}})$. P believes the contents of the message are recognizable.

Applying I1 we obtain $P \mid\equiv S \mid\sim (N_p, Q, K, \{K, P\}_{K_{qs}})$. P believes the message originated from S.

Applying J2 we obtain $P \mid\equiv S \mid\equiv P \overset{K}{\leftrightarrow} Q$. P believes S believes that K is a good key for P and Q.

Applying J1 we obtain $P \mid\equiv P \overset{K}{\leftrightarrow} Q$. P believes that K is a suitable key for P and Q.

Message 3: the extension to the message, $S \mid\equiv P \overset{K}{\leftrightarrow} Q$, is valid.

Applying T1, T3 and P1 we obtain $Q \ni K$. Q possesses K.

None of our postulates enable us to further derive new useful beliefs or possessions from this message. In particular, we cannot derive the freshness of the message. Indeed, as far as Q is concerned it could very well be a replay of a message from previous runs. Unlike the BAN logic, our rules cannot even provide the conclusion that S has *once* sent the message. Q has no reason to know this is not a replay of a message he once sent himself, unless a principal can identify messages that were not originated by themselves. Finally, because Q is not sure that S conveyed the message at all, or that it is fresh, Q obviously can not convince himself about S's current beliefs and cannot make use of the extension to the message.

Message 4: applying T1, T3 and P1 we obtain $P \ni N_q$. P possesses N_q.

No further conclusions can be derived. Although K is used, P does not know who the originator of the message is, or that Q now possesses the key. The reason is that N_q, a random number generated by Q, is not recognizable to P. Thus there is not sufficient redundancy for P to be convinced of the genuineness of the decrypted content.

Message 5: none of our postulates can derive any useful belief or possession from that message. In particular Q cannot even be convinced that it was sent by P, much less draw any conclusions on P's state. Although the contents are accessible to Q who possess K, and the contents are recognizable to Q (postulate R1), Q is not convinced that he shares the key with P.

If we include a notion of *somebody else*, then we can derive that Q believes that somebody (not himself) who possesses K sent the message.

To conclude, our reasoning leads us to the conclusions that given the first three messages of the protocol, the last two messages attain nothing of use and can thus be eliminated without weakening the final position. However some modifications enable the derivation of the much improved final position that was originally intended by the protocol authors.

7.3 The Enhanced Needham-Schroeder Protocol

Recently, Needham and Schroeder suggested the following modification to their original protocol [7]:

1. $P \rightarrow Q$: P
2. $Q \rightarrow P$: $\{P, N_{q1}\}_{K_{qs}}$
3. $P \rightarrow S$: $P, Q, N_p, \{P, N_{q1}\}_{K_{qs}}$
4. $S \rightarrow P$: $\{N_p, Q, K, \{K, N_{q1}, P\}_{K_{qs}}\}_{K_{ps}}$
5. $P \rightarrow Q$: $\{K, N_{q1}, P\}_{K_{qs}}$
6. $Q \rightarrow P$: $\{N_q\}_K$
7. $P \rightarrow Q$: $\{N_q - 1\}_K$

Q believes his nonce N_{q1} to be fresh. The difference between the two versions is that the enhanced version begins with an exchange between P and Q. P can thus provide Q's nonce to the server, and S includes that nonce in his response, which P forwards in message 5.

Similar reasoning to that of P, after message 2 in the original protocol, leads us to derive that after message 5 we obtain $Q \mid\equiv P \overset{K}{\leftrightarrow} Q$. Q believes that K is a suitable key for P and Q. That enables us to add the above belief as an extension to message 6:

$P \triangleleft *\{*N_q\}_K \rightsquigarrow Q \mid\equiv P \overset{K}{\leftrightarrow} Q$

However, P cannot deduce any conclusions from that message as it does not contain anything recognizable to P, as noted in the previous section. P does however gain the possession of N_q (postulates T1, T3 and P1).

Message 7: this message in the modified protocol is identical to message 5 in the original one:

$Q \triangleleft *\{*F(N_q)\}_K \rightsquigarrow P \mid\equiv P \overset{K}{\leftrightarrow} Q$

As usual, we first check that the extension is valid. Indeed it is, as we concluded that P holds that belief when he sends message 7.

Applying F1 and F2 we obtain $Q \mid\equiv \sharp(\{F(N_q)\}_K)$

Applying R1 we obtain $Q \mid\equiv \phi(F(N_q))$.

Applying I1 we obtain $Q \mid\equiv P \mid\sim (\{F(N_q)\}_K)$

Applying J2 we obtain $Q \mid\equiv P \mid\equiv P \overset{K}{\leftrightarrow} Q$

7.4 A Further Modification

The modification required to attain the desired final position is to include a formula in message 6 that would be recognizable to P. We propose the following modified version of message 6:

6. $P \triangleleft *\{*N_q, Q\}_K \rightsquigarrow Q \mid\equiv P \overset{K}{\leftrightarrow} Q$

Recalling that P believes the identifier Q to be recognizable (initial assumptions), similar reasoning to that described above would lead us to conclude:

$P \mid\equiv Q \mid\equiv P \overset{K}{\leftrightarrow} Q$

Or, if P does not trust Q's competence or honesty,

$P \mid\equiv Q \ni K$

7.5 Analysis Summary

We summarize the difference in the final positions that the above protocols attain. The original protocol attains a rather weak position:

$P \ni K; \quad P \mid\equiv P \overset{K}{\leftrightarrow} Q$

$Q \ni K$

That is, P possesses and believes a mutual secret, Q possesses the secret but cannot believe it. Neither believe anything about the other.

The extended protocol attains a stronger position even if P and Q do not trust each other's competence or honesty.

$$P \ni K; \quad P \mid\equiv P \overset{K}{\leftrightarrow} Q$$

$$Q \ni K; \quad Q \mid\equiv P \overset{K}{\leftrightarrow} Q; \quad Q \mid\equiv P \ni K$$

That is, P and Q possess the key and believe in it and Q believes that P possesses it. If in addition Q trusts P's competence and honesty, our reasoning process allows us to conclude that the extended protocol attains:

$$P \ni K; \quad P \mid\equiv P \overset{K}{\leftrightarrow} Q$$

$$Q \ni K; \quad Q \mid\equiv P \overset{K}{\leftrightarrow} Q; \quad Q \mid\equiv P \ni K;$$

$$Q \mid\equiv P \mid\equiv P \overset{K}{\leftrightarrow} Q$$

That is, P and Q possess the key and believe in it, and Q believes that P possesses it and believes in it. Finally, our modified version attains:

$$P \ni K; \quad P \mid\equiv P \overset{K}{\leftrightarrow} Q; \quad P \mid\equiv Q \ni K;$$

$$P \mid\equiv Q \mid\equiv P \overset{K}{\leftrightarrow} Q$$

$$Q \ni K; \quad Q \mid\equiv P \overset{K}{\leftrightarrow} Q; \quad Q \mid\equiv P \ni K;$$

$$Q \mid\equiv P \mid\equiv P \overset{K}{\leftrightarrow} Q$$

That is, P and Q possess the key and believe in it and each believes that the other possesses it and believes in it, assuming they trust each other. We note that the Yahalom protocol [1] can attain the same final position with five rather than seven messages.

8 Conclusions

We presented a new approach to reasoning about cryptographic protocols. Our work, which was inspired by the BAN logic work, offers significant advantages. Some of the main differences between the two approaches are summarized below.

The notion of *possession* incorporated in our approach assumes that principals can include in messages data they do not *believe* in, but merely possess. This also enables us to derive conclusions such as "Q possesses the shared key" in the example in the previous section. The BAN approach, on the other hand, requires all principals to include in messages only formulae in which they believe, and only allows the derivation of stronger conclusions of the kind "Q believes in the shared key", if they hold.

Our approach places a strong emphasis on the separation between the content and the meaning of messages. This can increase consistency in the analysis and, more importantly, introduce the ability to reason at more than one level. The final position in a given run will depend on the level of mutual trust of the specific principals participating in that run.

The notion of *recognizability* is an important one. With this notion, it is possible to express the ability of a recipient to identify the messages he expects. The BAN logic assumes that the encryption component always provides enough redundancy, which may not always be necessary. In fact, it is sometimes rather undesirable [4].

The *not-originated-here* notion allows us to determine that certain messages are not replays of a recipient's own previous messages in a session. BAN assumes that the encryption always provides information that identifies the sender. This can be achieved by, for example, embedding a sender's unique (and perhaps permanent) identifier in every message. We believe that such a property should not always be required, and in any case needs to be specified explicitly at the protocol description level.

The many additional constructs and rules we have introduced extend the scope of the protocols that can be analyzed. The new approach requires less assumptions and can thus be seen as more general. Also, it is not limited to authentication protocols, and can be used to analyze, for example, some cryptographic protocols that make use of one-way functions.

Like BAN, we choose not to include elaborate temporal notions or any form of negation. This helps to simplify the analysis process, while being sufficient for the derivation of important characteristics of cryptographic protocols. Extending that model is a subject of current investigation.

Finally we note that beliefs are not associated with any guarantee of truth with respect to the real world state they represent. It seems that beliefs are an appropriate notion to use in an environment of potential mutual suspicion such as the one we assume.

9 Acknowledgments

We would like to thank Martin Abadi and Mike Burrows of DEC Systems Research Center, Bill Harbison, Mark Lomas, and David Wheeler of the University of Cambridge Computer Laboratory, and the referees for valuable comments. Susanna Wills helped to improve the presentation.

References

[1] M. Burrows, M. Abadi, and R. Needham, "A Logic of Authentication", in Proceedings of the 12th ACM Symposium on Operating Systems Principles, Litchfield Park, Arizona, December, 1989. Published as *ACM Operating System Review*, Vol.23, No.5, pp.1-13, December, 1989. A fuller version was published as DEC System Research Center Report No.39, Palo Alto, California, February, 1989.

[2] J.Y. Halpern and Y. Moses, "Knowledge and Common Knowledge in a Distributed Environment", in Proceedings of the 3rd ACM Symposium on Principles of Distributed Computing, pp.50-61, Vancouver, British Columbia, August, 1984.

[3] C.A.R. Hoare, "An Axiomatic Basis for Computer Programming", *Communications of the ACM*, Vol.12, No.10, pp.576-580 and p.583, October, 1969.

[4] T.M.A. Lomas, L. Gong, J.H. Saltzer, and R.M. Needham, "Reducing Risks from Poorly Chosen Keys", in Proceedings of the 12th ACM Symposium on Operating Systems Principles, Litchfield Park, Arizona, December, 1989. Published as *ACM Operating System Review*, Vol.23, No.5, pp.14-18, December, 1989.

[5] R.C. Merkle, "One Way Hash Functions and DES", in *Advances of Cryptology*, Proceedings of Crypto '89, Santa Barbara, California, October, 1989.

[6] R.M. Needham and M.D. Schroeder, "Using Encryption For Authentication in Large Networks of Computers", *Communications of the ACM*, Vol.21, No.12, pp.993-999, December, 1978.

[7] R.M. Needham and M.D. Schroeder, "Authentication Revisited", *Operating Systems Review*, Vol.21, No.1, p.7, January, 1987.

[8] R.L. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-key Cryptosystems", *Communications of the ACM*, Vol.21, No.2, pp.120-126, February, 1978.

A Logical Postulates

In this appendix we list all the logical postulates and their description. A postulate that applies to formula X also applies to $*X$, though not necessarily vice versa.

We recall that these postulates can be supplemented with the following rationality rule. Informally, it states that our set of postulates can be expanded to permit reasoning about a principal's beliefs regarding the state of other principals. More precisely:

if $\frac{C1}{C2}$ is a postulate, then for any principal P, so is $\frac{P \mid\equiv C1}{P \mid\equiv C2}$.

A.1 Being-Told Rules

T1 $\quad \dfrac{P \lhd *X}{P \lhd X}$

Being told a "not-originated-here" formula is a special case of being told a formula.

T2 $\quad \dfrac{P \lhd (X,Y)}{P \lhd X}$

Being told a formula implies being told each of its concatenated components.

T3 $\quad \dfrac{P \lhd \{X\}_K,\ P \ni K}{P \lhd X}$

If a principal is told a formula encrypted with a key he possesses then he is considered to have also been told the decrypted contents of that formula.

T4 $\quad \dfrac{P \lhd \{X\}_{+K},\ P \ni -K}{P \lhd X}$

If a principal is told a formula encrypted with a public key and he possesses the corresponding private key then he is considered to have also been told the decrypted contents of that formula.

T5 $\quad \dfrac{P \lhd F(X,Y),\ P \ni X}{P \lhd Y}$

If a principal is told the result of a function F, and if he possesses one of the two arguments, then he is considered to have been told the other argument as well (see definition of F in section 3.1).

For RSA, or any public-key system with the property $\{\{X\}_{-K}\}_{+K} = X$, the following postulate holds:

T6 $\quad \dfrac{P \lhd \{X\}_{-K},\ P \ni +K}{P \lhd X}$

If a principal is told a formula encrypted with a private key and he possesses the corresponding public key then he is considered to have also been told the decrypted contents of that formula.

A.2 Possession Rules

P1 $\quad \dfrac{P \lhd X}{P \ni X}$

A principal is capable of possessing anything he is told.

P2 $\quad \dfrac{P \ni X,\ P \ni Y}{P \ni (X,Y),\ P \ni F(X,Y)}$

If a principal possesses two formulae then he is capable of possessing the formula constructed by concatenating the two formulae, as well as a computationally feasible function F of them.

P3 $\quad \dfrac{P \ni (X,Y)}{P \ni X}$

If a principal possesses a formula then he is capable of possessing any one of the concatenated components of that formula.

P4 $\quad \dfrac{P \ni X}{P \ni H(X)}$

If a principal possesses a formula then he is capable of possessing a one-way computationally feasible function of that formula.

P5 $$\frac{P \ni F(X,Y),\ P \ni X}{P \ni Y}$$

If a principal possesses a function F (as defined in section 3.1) of two formulae, and if he possesses one of these formula, then he is capable of possessing the other formula as well.

P6 $$\frac{P \ni K,\ P \ni X}{P \ni \{X\}_K,\ P \ni \{X\}_K^{-1}}$$

If a principal possesses a formula and a key then he is capable of possessing both the encryption and the decryption of the formula with the key.

P7 $$\frac{P \ni +K,\ P \ni X}{P \ni \{X\}_{+K}}$$

If a principal possesses a formula and a public key then he is capable of possessing the encryption of that formula with the key.

P8 $$\frac{P \ni -K,\ P \ni X}{P \ni \{X\}_{-K}}$$

If a principal possesses a formula and a private key then he is capable of possessing the decryption of that formula with the key.

A.3 Freshness Rules

For convenience, we use $P \mid\equiv \sharp(X,Y)$ to denote $P \mid\equiv \sharp(X)$ or $P \mid\equiv \sharp(Y)$.

F1 $$\frac{P \mid\equiv \sharp(X)}{P \mid\equiv \sharp(X,Y),\ P \mid\equiv \sharp(F(X))}$$

If P believes a formula X is fresh, then he is entitled to believe that any formula of which X is a component is fresh, and that a computationally feasible one-to-one function F of X is fresh.

F2 $$\frac{P \mid\equiv \sharp(X),\ P \ni K}{P \mid\equiv \sharp(\{X\}_K),\ P \mid\equiv \sharp(\{X\}_K^{-1})}$$

If P believes a formula X is fresh and possesses a key, then P is entitled to believe that the encryption and the decryption of X with the key are fresh.

F3 $$\frac{P \mid\equiv \sharp(X),\ P \ni +K}{P \mid\equiv \sharp(\{X\}_{+K})}$$

If P believes a formula X is fresh and possesses a public key, then P is entitled to believe that the encryption of X with that key is fresh as well.

F4 $$\frac{P \mid\equiv \sharp(X),\ P \ni -K}{P \mid\equiv \sharp(\{X\}_{-K})}$$

If P believes a formula X is fresh and possesses a private key, then P is entitled to believe that the decryption of X with that key is fresh as well.

F5 $$\frac{P \mid\equiv \sharp(+K)}{P \mid\equiv \sharp(-K)}$$

If P believes that a public key is fresh, then he is entitled to believe that the corresponding private key is fresh as well.

F6 $$\frac{P \mid\equiv \sharp(-K)}{P \mid\equiv \sharp(+K)}$$

If P believes that a private key is fresh, then he is entitled to believe that the corresponding public key is fresh as well.

F7 $$\frac{P \mid\equiv \phi(X),\ P \mid\equiv \sharp(K),\ P \ni K}{P \mid\equiv \sharp(\{X\}_K),\ P \mid\equiv \sharp(\{X\}_K^{-1})}$$

If P believes that a formula X is recognizable and P possesses a key K and believes it is fresh, then P is entitled to believe that the encryption and the decryption of X with K are fresh.

F8 $$\frac{P \mid\equiv \phi(X),\ P \mid\equiv \sharp(+K),\ P \ni +K}{P \mid\equiv \sharp(\{X\}_{+K})}$$

If P believes that a formula X is recognizable, and P possesses a public key and believes it is fresh then P is entitled to believe that the encryption of X with that public key is fresh.

F9 $$\frac{P \mid\equiv \phi(X),\ P \mid\equiv \sharp(-K),\ P \ni -K}{P \mid\equiv \sharp(\{X\}_{-K})}$$

If P believes that a formula X is recognizable and P possesses a private key and believes it is fresh, then P is entitled to believe that the decryption of X with that private key is fresh.

F10 $$\frac{P \mid\equiv \sharp(X),\ P \ni X}{P \mid\equiv \sharp(H(X))}$$

If P believes a formula X is fresh and he also possesses X, then he is entitled to believe that a one-way function of X is fresh.

F11 $$\frac{P \mid\equiv \sharp(H(X)),\ P \ni H(X)}{P \mid\equiv \sharp(X)}$$

If P possesses $H(X)$ and believes it to be fresh, then he is entitled to believe that X is fresh.

A.4 Recognizability Rules

R1
$$\frac{P \mid\equiv \phi(X)}{P \mid\equiv \phi(X,Y),\ P \mid\equiv \phi(F(X))}$$

If P believes a formula X is recognizable, then he is entitled to believe that any formula of which X is a component is recognizable, and that a computationally feasible function F of X is recognizable.

R2
$$\frac{P \mid\equiv \phi(X),\ P \ni K}{P \mid\equiv \phi(\{X\}_K),\ P \mid\equiv \phi(\{X\}_K^{-1})}$$

If P believes a formula X is recognizable and P possesses a key K, then P is entitled to believe that the encryption and the decryption of X with K are recognizable.

R3
$$\frac{P \mid\equiv \phi(X),\ P \ni +K}{P \mid\equiv \phi(\{X\}_{+K})}$$

If P believes a formula X is recognizable and P possesses a public key , then P is entitled to believe that the encryption of X with that key is recognizable.

R4
$$\frac{P \mid\equiv \phi(X),\ P \ni -K}{P \mid\equiv \phi(\{X\}_{-K})}$$

If P believes a formula X is recognizable and P possesses a private key, then P is entitled to believe that the decryption of X with that key is recognizable.

R5
$$\frac{P \mid\equiv \phi(X),\ P \ni X}{P \mid\equiv \phi(H(X))}$$

If P believes a formula X is recognizable and he also possesses X, then he is entitled to believe that a one-way function of X is recognizable.

R6
$$\frac{P \ni H(X)}{P \mid\equiv \phi(X)}$$

If P possesses formula $H(X)$ then he is entitled to believe that X is recognizable.

A.5 Message Interpretation Rules

A secret S used for identification purposes is denoted as $< S >$. This way it can be distinguished when other secrets are transmitted as data in the same message.

I1
$$\frac{P \triangleleft *\{X\}_K,\ P \ni K,\ P \mid\equiv P \overset{K}{\leftrightarrow} Q,\ P \mid\equiv \phi(X),\ P \mid\equiv \sharp(X,K)}{P \mid\equiv Q \mid\sim X,\ P \mid\equiv Q \mid\sim \{X\}_K,\ P \mid\equiv Q \ni K}$$

Suppose that for principal P all of the following conditions hold: (1) P receives a formula consisting of a X encrypted with key K and marked with a not-originated-here mark; (2) P possesses K; (3) P believes K is a suitable secret for himself and Q; (4) P believes formula X is recognizable; (5) P believes that K is fresh or that X is fresh.

Then P is entitled to believe that (1) Q once conveyed X; (2) Q once conveyed the formula X encrypted with K; (3) Q possesses K.

I2
$$\frac{P \triangleleft *\{X, < S >\}_{+K},\ P \ni (-K, S),\ P \mid\equiv^{+K}\!\!\mapsto P,\ P \mid\equiv P \overset{S}{\leftrightarrow} Q,\ P \mid\equiv \phi(X,S),\ P \mid\equiv \sharp(X,S,+K)}{P \mid\equiv Q \mid\sim (X, < S >),\ P \mid\equiv Q \mid\sim \{X, < S >\}_{+K},\ P \mid\equiv Q \ni +K}$$

Suppose that for principal P all of the following conditions hold: (1) P receives a formula consisting of X concatenated with S, encrypted with a public key and marked with a not-originated-here mark; (2) P possesses S and the corresponding private key; (3) P believes the public key is his own; (4) P believes S is a suitable secret for himself and Q; (5) P believes that X concatenated with S is recognizable; (6) P believes that at least one of $S, X, or + K$ is fresh.

Then P is entitled to believe that (1) Q once conveyed the formula X concatenated with S; (2) Q once conveyed the formula X concatenated with S and encrypted with the public key; (3) Q possesses the public key.

I3
$$\frac{P \triangleleft *H(X, < S >),\ P \ni (X,S),\ P \mid\equiv P \overset{S}{\leftrightarrow} Q,\ P \mid\equiv \sharp(X,S)}{P \mid\equiv Q \mid\sim (X, < S >),\ P \mid\equiv Q \mid\sim H(X, < S >)}$$

Suppose that for principal P all of the following conditions hold: (1) P receives a formula consisting of a one-way function of X and S marked with a not-originated-here mark; (2) P possesses S and X; (3) P believes S is a suitable secret for himself and Q; (4) P believes that either S or X is fresh.

Then P is entitled to believe that (1) Q once conveyed the formula X concatenated with S; (2) Q once conveyed the one-way function of X concatenated with S.

To understand I3 on the use of one-way functions, it helps to compare it with the previous two. Also, recall that according to our definition $Q \mid\sim (X,S)$ does not mean that Q necessarily included and transmitted X and S explicitly in any message, but rather that he conveyed X and S.

Postulates I4 and I5 hold only for RSA or other public-key schemes with a similar property to the one described in section 3.

I4
$$\frac{P \triangleleft \{X\}_{-K},\ P \ni +K,\ P \mid\equiv^{+K}\!\!\mapsto Q,\ P \mid\equiv \phi(X)}{P \mid\equiv Q \mid\sim X,\ P \mid\equiv Q \mid\sim \{X\}_{-K}}$$

Suppose for principal P, all of the following conditions hold: (1) P receives a formula consisting of X encrypted with a private key; (2) P possesses the corresponding public key; (3) P believes that public key is Q's; (4) P believes X is recognizable.

Then P is entitled to believe that (1) Q once conveyed the formula X; (2) Q once conveyed the formula consisting of X encrypted with the private key.

I5

$$\frac{P \triangleleft \{X\}_{-K},\ P \ni +K,\ P \mid\equiv \overset{+K}{\mapsto} Q,\ P \mid\equiv \phi(X),\ P \mid\equiv \sharp(X, +K)}{P \mid\equiv Q \ni (-K, X)}$$

Suppose for principal P, all of the following conditions hold: (1) P receives a formula consisting of X encrypted with a private key; (2) P possesses the corresponding public key; (3) P believes that public key is Q's; (4) P believes X is recognizable; (5) P believes that either X or the public key is fresh.

Then P is entitled to believe that Q possesses the string consisting of the concatenation of X and the private key.

I6 $$\frac{P \mid\equiv Q \mid\sim X,\ P \mid\equiv \sharp(X)}{P \mid\equiv Q \ni X}$$

If P believes that Q once conveyed formula X and P believes that X is fresh, then P is entitled to believe that Q possesses X.

I7 $$\frac{P \mid\equiv Q \mid\sim (X, Y)}{P \mid\equiv Q \mid\sim X}$$

If P believes that Q once conveyed the formula consisting of the concatenation of X and Y, then P is entitled to believe that Q once conveyed X.

A.6 Jurisdiction Rules

J1 $$\frac{P \mid\equiv Q \mid\Longrightarrow C,\ P \mid\equiv Q \mid\equiv C}{P \mid\equiv C}$$

If P believes that Q is an authority on some statement C and that Q believes in C, then P ought to believe in C as well.

J2 $$\frac{P \mid\equiv Q \mid\Longrightarrow Q \mid\equiv *,\ P \mid\equiv Q \mid\sim (X \rightsquigarrow C),\ P \mid\equiv \sharp(X)}{P \mid\equiv Q \mid\equiv C}$$

If P believes that Q is honest and competent, and P receives a message $X \rightsquigarrow C$ which he believes Q conveyed, then P ought to believe that Q really believes C.

J3 $$\frac{P \mid\equiv Q \mid\Longrightarrow Q \mid\equiv *,\ P \mid\equiv Q \mid\equiv Q \mid\equiv C}{P \mid\equiv Q \mid\equiv C}$$

If P believes that Q is honest and competent, and P believes that Q believes that Q believes in C, then P ought to believe that Q believes in C.

B "Never-Originated-Here" Messages

The "not-originated-here" notion in section 3.3 refers to messages in the current session. Some principals may be able to identify that certain messages were not originated by themselves in current or previous sessions. Suppose we use $P \mid\equiv \otimes(P)$ to denote that P believes he can identify such messages, then we can add the following postulates.

I1′ $$\frac{P \triangleleft \{X\}_K,\ P \ni K,\ P \mid\equiv P \overset{K}{\leftrightarrow} Q,\ P \mid\equiv \phi(X),\ P \mid\equiv \otimes(P)}{P \mid\equiv Q \mid\sim X,\ P \mid\equiv Q \mid\sim \{X\}_K}$$

Suppose that for principal P all of the following conditions hold: (1) P receives a formula consisting of X encrypted with key K and marked with a not-originated-here mark; (2) P possesses K; (3) P believes K is a suitable secret for himself and Q; (4) P believes formula X is recognizable; (5) P believes that he can identify that a message was not originated by himself.

Then P is entitled to believe that Q once conveyed X and Q once conveyed X encrypted with K.

I2′

$$\frac{P \triangleleft \{X, < S >\}_{+K}, P \ni (S, -K),\ P \mid\equiv \overset{+K}{\mapsto} P,\ P \mid\equiv P \overset{S}{\leftrightarrow} Q,\ P \mid\equiv \phi(X, S), P \mid\equiv \otimes(P)}{P \mid\equiv Q \mid\sim (X, < S >),\ P \mid\equiv Q \mid\sim \{X, < S >\}_{+K}}$$

Suppose that for principal P all of the following conditions hold: (1) P receives a formula consisting of X concatenated with S, encrypted with a public key and marked with a not-originated-here mark; (2) P possesses S and the corresponding private key; (3) P believes the public key is his own; (4) P believes S is a suitable secret for himself and Q; (5) P believes that X concatenated with X is recognizable; (6) P believes he can identify that a message was not originated by himself.

Then P is entitled to believe that Q once conveyed the formula X concatenated with S, and Q once conveyed the formula X concatenated with S encrypted with the public key.

I3′

$$\frac{P \triangleleft H(X, < S >),\ P \ni (X, S),\ P \mid\equiv P \overset{S}{\leftrightarrow} Q,\ P \mid\equiv \phi(X, S),\ P \mid\equiv \otimes(P)}{P \mid\equiv Q \mid\sim (X, < S >),\ P \mid\equiv Q \mid\sim H(X, < S >)}$$

Suppose that for principal P all of the following conditions hold: (1) P receives a formula consisting of a one-way function of X and S marked with a not-originated-here mark; (2) P possesses S and X; (3) P believes S is a suitable secret for himself and Q; (4) P believes that S concatenated with X is recognizable; (5) P believes he can identify that a message was not originated by himself.

Then P is entitled to believe that Q once conveyed the formula X concatenated with S, and Q once conveyed the one-way function of X concatenated with S.

It is possible, and perhaps worthwhile, to refine the notion by associating such an assumption with particular formulae. Thus for example in I1$'$ above we would replace $P \mid\equiv \otimes(P)$ with $P \mid\equiv \otimes(\{X\}_K)$.

C Semantics

The semantics are similar to those in BAN [1]. The semantics are "operational". Principals develop new beliefs and accumulate possessions by applying inference rules to their present beliefs, possessions, and received messages.

- The local state of a principal P consists of two sets. A set of formulae $\mathcal{P}_P$ and a set of statements $\mathcal{B}_P$. Intuitively, $\mathcal{P}_P$ is the set of formulae the principal possesses and $\mathcal{B}_P$ is the set of beliefs of the principal. The sets have some closure properties, as reflected in the inference rules.

- A global state is a tuple containing the local states of all principals. Suppose s is a global state then s_P is the local state of P in s and $\mathcal{P}_P(s)$ and $\mathcal{B}_P(s)$ are the corresponding sets of possession and beliefs. The satisfaction relation between global states and statements is: $P \mid\equiv$ C holds in a state s if C $\in \mathcal{B}_P(s)$ and $P \ni X$ holds if $X \in \mathcal{P}_P(s)$. A set of statements holds in a given state if each of its members holds.

- A run is a finite sequence of states $s_0, \ldots, s_n$ where $\mathcal{B}_P(s_i) \subseteq \mathcal{B}_P(s_{i+1})$, and $\mathcal{P}_P(s_i) \subseteq \mathcal{P}_P(s_{i+1})$ for all $i \leq (n-1)$. That is, the belief set and the possess set cannot decrease during a run.

- A protocol is a finite sequence of n expressions of the form:

  $(P_1 \rightarrow Q_1 : X_1 \rightsquigarrow C_1), \ldots, (P_n \rightarrow Q_n : X_n \rightsquigarrow C_n)$

- In a run of a protocol all messages of the protocol are communicated. It is a run of length $n+1$ where

  $X_i \in \mathcal{P}_{Qi}(s_i) \cap \mathcal{P}_{Pi}(s_{i-1})$ and $C_n \in \mathcal{B}_{Pi}(s_{i-1})$, for all $i \leq n$.

- An annotation for the protocol holds in a run of the protocol if all of the statements in the annotation hold in the corresponding states. An annotation is valid if it holds in all runs of the protocol where the first set of the annotation – the assumptions – holds.