Completeness Theorems for the Abadi-Rogaway Language of Encrypted Expressions

Daniele Micciancio  Bogdan Warinschi
Computer Science Department, University of California, San Diego 9500 Gilman Drive, CA 92093 {daniele,bogdan}@cs.ucsd.edu

Abstract

We show that the Abadi-Rogaway logic of indistinguishability for cryptographic expressions is not complete, giving a natural example of encryption function (satisfying the standard notion of security) and a pair of expressions, such that the distributions associated to the two expressions are computationally indistinguishable, but equality cannot be proved within the Abadi-Rogaway logic. Then, we show that if a stronger, yet natural, notion of security for encryption is used (namely, that of authenticated encryption), then the Abadi-Rogaway logic is both sound and complete. In addition, we consider a refinement of the Abadi-Rogaway logic that overcomes certain limitations of the original proposal, allowing for encryption functions that do not hide the length of the message being sent. Both the soundness theorem of Abadi and Rogaway, and our completeness result for authenticated encryption easily extend to this more realistic notion of secrecy.

1 Introduction

In the last two decades, the logic and complexity theory communities have developed two related, though quite different, approaches to the theoretical modeling of security. Each approach has its own benefits and shortcomings. The logic approach, mostly syntactical, is primarily intended to make the task of writing and verifying cryptographic protocols more manageable. The major problem of this approach is that, without a satisfactory model theory behind, it is not quite clear how to interpret the security statements proved within these frameworks.

On the other hand, formalizing and clarifying the meaning of common security notions like secrecy, authenticity, etc., has been one of the primary goals and major achievements of the complexity based approach to cryptography. The stronger and more comprehensive adversarial model used by this approach has gained wide acceptance in the cryptographic community as the one providing the “right” notion of security. However, the relatively unstructured way in which security is formalized in this framework, makes proving the security of even moderately complex protocols quite a formidable task.

Recently, several attempts have been made to bridge the gap between these two worlds [10, 7, 1, 9]. Still, most of these approaches do not retain the simplicity of the original logic formulations, or do not capture the complexity theoretic notion of security in its full power. In a recent paper [2], Abadi and Rogaway defined a language of cryptographic expressions, and gave a simple semantics such that if two expressions have the same semantics then they are computational indistinguishable. The language Exp considered in [2] consists of all expressions that can be built from basic messages and cryptographic keys, using concatenation and encryption operation, according to the grammar:

\[
M ::= \ K \quad \text{key (for } K \in \text{Keys)} \\
\ i \quad \text{bit (for } i \in \text{Bool)} \\
\ (M, M) \quad \text{pair} \\
\ \{M\}_K \quad \text{encryption (for } K \in \text{Keys)}
\]

(1)

Abadi and Rogaway also define a simple (and easily computable) semantics that associates to each expression \( E \) an element of the language \( \text{Pattern} \) defined
by the following grammar:

\[
P ::= K \quad \text{key} \ (\text{for } K \in \text{Keys}) \\
i \quad \text{bit} \ (\text{for } i \in \text{Bool}) \\
(P, Q) \quad \text{pair} \\
\{P\}_K \quad \text{encryption} \ (\text{for } K \in \text{Keys}) \\
\square \quad \text{undecryptable}
\]

(2)

(see [2] for a definition of the pattern evaluation function). The main result of [2] is that if two expressions have the same pattern, then the probability distributions naturally associated to the two expressions (when keys are chosen independently at random, and encryption scheme satisfies a strong notion of security) are computationally indistinguishable. The standard notion of security for encryption functions [6] is that for any two messages \( m_0 \) and \( m_1 \), and randomly chosen key \( k \), the two ciphertexts \( E_k(m_0) \) and \( E_k(m_1) \) are computationally indistinguishable. The intuition behind this definition is that not only recovering the plaintext \( m \) from \( E_k(m) \) must be hard, but also computing any partial information about \( m \) should be unfeasible. Similarly, the indistinguishability of the probability distributions associated to two expressions with the same pattern, captures the computational notion that the encoded expressions do not leak any partial information other than what was intended by the protocol designer, as specified by the pattern.

The cryptographic language (1) is clearly limited as it allows to model only the simplest cryptographic protocols where one party sends a sequence of (complex) messages to another one. However, [2] is a nice example of a simple language with a clean semantics which is sound according to the standard notion of security of complexity based cryptography. So, [2] is an interesting first step toward languages that can be used to model more complex cryptographic protocols. In this paper we give a closer look at the semantics of [2], and consider the following two important issues:

- Abadi and Rogaway consider encryption functions that hide every possible information about the message. This is not a realistic assumption, because in order to allow unique decryption, the ciphertext must be at least as long as the plaintext. In particular, the length of the message being encrypted cannot be completely hidden without making the scheme totally impractical.
- Another important issue, is that of completeness, i.e., if the probability associated to two expressions are indistinguishable for every secure encryption function, is it true that the two expressions have the same pattern? In other words, we ask whether the pattern associated to an expression fully characterizes the information recoverable by an adversary, or it is possible for two expressions to be equivalent in the computational setting, and still equivalence cannot be proved using the Abadi-Rogaway pattern semantics.

Regarding the first issue, as clearly stated in [2], this is only a simplifying assumption to make the exposition more readable, and their semantics and soundness theorem can be extended to more realistic encryption schemes that hide all information about the plaintext other than its length.

Regarding the second point, [2] already gives an example to illustrate that their semantics is not complete. However, the example is quite artificial because it involves an encryption function with restricted message space. If we consider general purpose encryption schemes (i.e., schemes that can be used to encrypt arbitrary messages), then the example of [2] breaks down, leaving the question of completeness open. In this paper we show that the Abadi-Rogaway pattern semantics is not complete, even for general purpose encryption schemes satisfying the standard notion of security. The example is presented in Section 2. Then, in Section 3, we consider a stronger notions of security for encryption schemes. Namely, we consider authenticated encryption schemes, i.e., encryption schemes providing both secrecy and authenticity. This is a stronger, yet natural and relatively standard, notion of security and it has been the subject of many recent research papers in the cryptographic community [5, 3, 8]. Interestingly, we can show that if authenticated encryption is used, then the Abadi-Rogaway pattern semantics is both sound and complete. This is the main technical contribution of this note, and it is proved in Section 3.

Finally, in Section 4, we extend the pattern semantics of [2] to deal with encryption schemes that do not hide the length of the message, and observe that both the soundness and completeness results hold for these more realistic encryption schemes as well.

2 Incompleteness of the Abadi-Rogaway semantics

Abadi and Rogaway define a function \( \text{rec}(\cdot) \) that associates to every expression \( E \) the set of keys recoverable from \( E \) (see [2] for a formal definition). For example

\[
\text{rec}((\{K_1\}, K_2, \{0\}_{K_2}, K_1, \{K_2\}, K_1)) = \{K_1, K_2\}
\]

The function \( p : \text{Exp} \times \mathcal{P}(\text{Keys}) \to \text{Pattern} \), also
introduced in [2], associates to every pair \((E, T) \in \text{Exp} \times \mathcal{P}(\text{Keys})\) the pattern obtained by replacing encryptions under keys outside the set \(T\) with the \(\square\) symbol. If \(E\) is as in the previous example,
\[
 p(E, \{K_2\}) = (\square, \{0\} K_1, K_1, \square)
\]

Such a pattern models the partial information an adversary can get from \(E\) having the knowledge of the keys in \(T\). Finally, the pattern associated to an expression is defined as \(p(E, \text{rec}(E))\) and two cryptographic expressions \(E_1, E_2\) are deemed equivalent if their associated patterns are equal up to renaming of the keys. We write \(E_1 \cong E_2\) to denote this equivalence.

From a computational perspective, a symmetric encryption scheme \(\Pi\) is given by three algorithms: \(\Pi\) the key generation algorithm takes as input a security parameter \(\eta\) and outputs a key \(k\) in some domain \(\text{Key}\). The encryption algorithm \(\mathcal{E}\) takes as input a key \(k\) and a plaintext \(m\) and outputs a ciphertext \(\mathcal{E}_k(m)\). Finally, the decryption algorithm \(\mathcal{D}\), takes as input a key \(k\) and a string \(c\) and outputs the underlying plaintext or \(\perp\) (used to indicate decryption failures.) It is required that for any string \(m\) and any key \(k\),
\[
 \mathcal{D}_k(\mathcal{E}_k(m)) = m.
\]

For an encryption scheme \(\Pi\), fixing a security parameter \(\eta\) induces a natural mapping \([\cdot]_{\Pi(\eta)}\) from expressions to probability distributions on strings. For an expression \(E\), random keys are assigned to the key names appearing in \(E\). Then, the distribution is inductively defined following the structure of \(E\). Assuming what is called in [2] a type-0 secure encryption scheme \(\Pi\) (namely an encryption scheme that hides all partial information on the key, and all partial information on the plaintext) the following connection can be shown between the formal and the computational approach:

**Theorem 2.1** Let \(E_0\) and \(E_1\) be acyclic expressions and let \(\Pi\) be a type-0-secure encryption scheme. Suppose that \(E_0 \cong E_1\), then \([E_0]_{\Pi(\eta)} \approx [E_1]_{\Pi(\eta)}\).

Next we provide a counterexample for the converse of the theorem. Although in doing so we drop the requirement that \(\Pi\) hides the length of the plaintext, we shed some light on the properties that might be needed in order to obtain the converse result.

Consider the following encryption scheme based on a family \(\{F_j\}_{j \in I}\) of pseudorandom functions. The key generation algorithm simply outputs a random index \(k \in I\). The encryption of a message \(m\) is \(\mathcal{E}_k(m) = r \| F_k(r) \oplus m\), where \(r\) is some generated randomness. The decryption of a message \((r|c)\) is simply \(F_k(r) \oplus c\). Consider the expressions
\[
 E_1 = (\{K_3\} K_1, K_1) \quad \text{and} \quad E_2 = (\{K_3\} K_1, K_2).
\]

Then, while the patterns associated:
\[
 \text{pattern}(E_1) = (\{K_3\} K_1, K_1) \\
 \text{pattern}(E_2) = (\square, K_2),
\]

are clearly different, the distributions of \((r_1 \| F_k(r_1) \oplus k_3, k_1)\) and \((r_2 \| F_k(r_2) \oplus k_3, k_2)\) are computationally indistinguishable. In fact, they are identical when \(r_1, r_2, k_1, k_2, k_3\) are picked at random. This is related to the fact that knowledge of a key \(k\) does not imply the ability of distinguishing valid decryptions under \(k\) from invalid ones.

**3 A completeness theorem for authenticated encryption**

Next we consider secure authenticated encryption schemes. Informally, such schemes provide a mechanism for checking the validity of a ciphertext. The authenticity of the plaintext is ensured, if no efficient adversary can create a valid ciphertext except with negligible probability. The formal definition below, follows [3, 5].

**Definition 3.1** [Integrity of an Authenticated Encryption Scheme] Let \(\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})\) be a symmetric encryption scheme and \(\eta \in \mathbb{N}\) a security parameter. Consider the following experiment involving an algorithm \(A\) with access to an encryption oracle:
\[
 \text{Exp}_{\Pi,A}^{\text{int-pxt}}(\eta) = \begin{cases} 
 K &\in \mathcal{K}(\eta) \\
 C &\in \mathcal{A}^{|K(\cdot)(\eta)} \\
 \text{If } \mathcal{D}_K(C) \neq \perp \text{ and } C \text{ was never} &\text{a response of } \mathcal{E}_K(\cdot) \text{ then return 0} \\
 \text{else return 1} &
\end{cases}
\]

Define the advantage of the algorithm \(A\) as:
\[
 \text{Adv}_{\Pi,A}^{\text{int-pxt}}(\eta) = \Pr[\text{Exp}_{\Pi,A}^{\text{int-pxt}}(\eta) = 1].
\]

The encryption scheme \(\Pi\) is said to be INT-PTXT secure if the advantage function \(\text{Adv}_{\Pi,A}^{\text{int-pxt}}(\cdot)\) is negligible for any adversary \(A\) whose time-complexity is polynomial in \(\eta\).

As a consequence, for a INT-PTXT secure encryption scheme, decrypting with key \(k_1\) the encryption of some message under the key \(k_2\) fails with high probability. This is captured by the following lemma:

**Lemma 3.2** Let \(\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})\) be a INT-PTXT secure symmetric authenticated encryption scheme. Then for any \(x \in \text{String}\) and \(\eta \in \mathbb{N}\),
\[
 \Pr_{k_1, k_2 \in \mathcal{K}(\eta), x}[\mathcal{D}_{k_1}(\mathcal{E}_{k_2}(x))] \leq \nu_0(\eta), \text{ for some negligible function } \nu_0(\cdot).
\]
Since the proof of the lemma exhibits an adversary which does not make use of the encryption oracle, one could consider weaker security notions for $\Pi$. We remark that although most of our statements use encryption schemes that are \textsc{INT-PTXT} secure, the same results would hold for the weaker requirement given by the conclusion of Lemma 3.2.

### 3.1 The Completeness Theorem

We show that for authenticated encryption, the pattern semantics agrees with the computational semantics. More precisely we will prove the following theorem:

**Theorem 3.3** Let $\Pi$ be an \textsc{INT-PTXT} secure encryption scheme. Then if $E_0$ and $E_1$ are acyclic expressions such that $E_0 \not\equiv E_1$ then $[E_0]_{\Pi[\eta]} \not\equiv [E_1]_{\Pi[\eta]}$.

The proof is based on the intuition that from a sample from $[E]\Pi[\eta]$ one can recover the pattern of $E$. Next we define some auxiliary functions that will prove to be useful in reasoning about the recovery algorithm.

The function $F_{kr} : \text{Exp} \times \mathcal{P}\left(\text{Keys}\right) \to \mathcal{P}\left(\text{Keys}\right)$ further defined by

- $F_{kr}(i, T) = T$;
- $F_{kr}(K, T) = \{K\} \cup T$;
- $F_{kr}(\{E_0, E_1\}, T) = F_{kr}(E_0, T) \cup F_{kr}(E_1, T)$;
- $F_{kr}(\{E\}_K, T) = T$, if $K \not\in T$;
- $F_{kr}(\{E\}_K, T) = F_{kr}(E, T)$, if $K \in T$;

computes the set $F_{kr}(E, T)$ containing the keys that can be recovered from $E$ having knowledge of the keys in $T$ going through $E$ only once. Using the auxiliary functions $G_i : \text{Exp} \to \mathcal{P}\left(\text{Keys}\right)$ defined by the relations:

- $G_0(E) = \emptyset$;
- $G_i(E) = F_{kr}(E, G_{i-1}(E))$

we define $F_{ikr}$, the iterative version of $F_{kr}$ as:

$$F_{ikr}(E) = \bigcup_i G_i(E).$$

The following equality holds: $\text{rec}(E) = F_{ikr}(E)$.

**Example.** If $E = (\{K_2\}_{K_1}, \{K_3\}_{K_2}, K_1)$ then $F_{kr}(E, \emptyset) = \{K_1\}$ and $F_{kr}(E, \{K_1\})$. According to the definition of $G_i$, $G_1(E) = \{K_1\}$ and $G_2(E) = \{K_1, K_2\}$. Also, $F_{ikr}(E) = \{K_1, K_2\}$.

The intuition that when given a sample from the distribution associated to an expression $E$ one can recover partial information (e.g. keys that are sent in clear), is captured by the following two functions. The key recovery function $kr : \text{String} \times \mathcal{P}\left(\text{Key}\right) \to \mathcal{P}\left(\text{Key}\right)$ and its iterative version $ikr : \text{String} \to \text{Key}$ are defined by the following two algorithms:

**Algorithm $kr(e, T)$**

- if $e = (i, \text{"boot"})$ output $T$;
- if $e = (k, \text{"key"})$ output $T \cup \{k\}$;
- if $e = (m, n)$, \text{"pair"} output $kr(m, T) \cup kr(n, T)$;
- if $e = (c, \text{"ciphertext"})$
  - if for exactly one $k \in T, D_k(c) \neq \perp$
  - then output $kr(D_k(c))$
  - else output $T$.

**Algorithm $ikr(e)$**

- $T_0 \leftarrow \emptyset$; $cont \leftarrow \text{true}$
  - while $cont$
    - $T_{i+1} \leftarrow kr(e, T_i)$
    - if $T_{i+1} = T_i$ then $cont \leftarrow \text{false}$
    - else $i \leftarrow i + 1$
  - output $T_i$

The relation between these two functions and their syntactic counterparts ($F_{kr}$ and $F_{ikr}$) is established by the following two lemmas (their proof as well as the proofs of other lemmas are omitted due to space constraints):

**Lemma 3.4** For any \textsc{INT-PTXT} secure authenticated encryption scheme $\Pi = (K, \mathcal{E}, \mathcal{D})$, any expression $E \in \text{Exp}$, any set of keys $T \subseteq \text{Keys}$ and any security parameter $\eta \in \mathbb{N}$, \Pr$[\text{kr}(e, \tau(T)) \neq \{r(F_{kr}(E, T))\}] \leq |E||T|\nu_0(\eta)$.

The probability is taken on $e$ $\in$ $[E]_{\Pi[\eta]}$ and $\nu_0(\cdot)$ is as in Lemma 3.2.

**Lemma 3.5** For any \textsc{INT-PTXT} secure authenticated encryption scheme $\Pi = (K, \mathcal{E}, \mathcal{D})$, any $E \in \text{Exp}$ and any security parameter $\eta \in \mathbb{N}$, \Pr$[\text{kr}(e) \neq \tau(F_{ikr}(E))] \leq |E|\nu_0(\eta)$.

In a similar manner we associate a computational counterpart to the function $p$ described in Section 2.

Let $\text{psp} : \text{String} \times \mathcal{P}(\text{Key}) \rightarrow \text{Pattern}$ be the function computed by the following algorithm:

**Algorithm $\text{psp}(e, tT)$**

- let $f : tT \rightarrow \text{Keys}$ be an injective function
  - if $e = (i, \text{"boot"})$ output $i$;
  - if $e = (k, \text{"key"})$ and $k \in tT$ then output $f(k)$;
  - if $e = (m, n)$, \text{"pair"} output $\langle\text{psp}(m, tT), \text{psp}(n, tT)\rangle$;
  - if $e = (c, \text{"ciphertext"})$
    - if for exactly one key $k \in tT, D_k(c) \neq \perp$
    - then output $\langle\text{psp}(D_k(c), tT)\rangle_{f(k)}$
    - else output $\perp$.

The following connection can be shown between $p$ and $\text{psp}$:

**Lemma 3.6** For any secure authenticated encryption scheme $\Pi = (K, \mathcal{E}, \mathcal{D})$, for any expression $E \in \text{Exp}$, any set of key names
$T \subseteq \text{Keys}$ and any security parameter $\eta$, 
$\Pr[p_{\text{psp}}(e_1, T) \neq \text{patternp}(E_i, T)] \leq |E| |T| \nu_0(\eta)$, 
where the probability is taken on $e \overset{R}{\leftarrow} [E]_n$ and $\nu_0(\cdot)$ is as in Lemma 3.2.

We can give now the proof of Theorem 3.3.

**Proof:**
Consider two arbitrary expressions $E_0, E_1 \in \text{Exp}$ such that $E_0 \neq E_1$. We will exhibit an algorithm that distinguishes between $[E_0]_n$ and $[E_1]_n$ with non-negligible probability:

Let $A$ be the following algorithm:

Algorithm $A(e)$
\[ tT \leftarrow \text{ikr}(e) \]
\[ \text{if } p_{\text{psp}}(e_1, tT) \equiv p(E_0, \text{rec}(E_0)) \text{ then output } 0 \]
\[ \text{else output } 1 \]

For $e_0 \overset{R}{\leftarrow} [E_0]_n$, denote by $A_t$ the event that $p_{\text{psp}}(e_i, tT) \equiv p(E_0, \text{rec}(E_0))$, for $i = 1, 2$. Then:
\[ \Pr[A_0] \geq \Pr[A_0 \land (tT = \tau(\text{rec}(E_0)))] = \Pr[A_0 \mid tT = \tau(\text{rec}(E_0))] \Pr[\text{ikr}(e) = \tau(\text{rec}(E_0))] \]

From Lemma 3.6 and Lemma 3.5 it follows that
\[ \Pr[A_0 \mid tT = \tau(\text{rec}(E_0))] \geq (1 - |E_0| |T| \nu_0(\eta)) \]

and
\[ \Pr[\text{ikr}(e) = \tau(\text{rec}(E_0))] \geq (1 - |E| |T| \nu_0(\eta)). \]

In a similar manner we can show that $\Pr[A_1] \geq (1 - 2|E_1| |T| \nu_0(\eta))$. Then:
\[ \Pr[A(e_0) = 0] = \Pr[A_0] \geq 1 - 2|E_0|^3 \nu_0(\eta). \]

(E_0 \neq E_1 implies $p(E_0, \text{rec}(E_0)) \neq p(E_1, \text{rec}(E_1))$). Therefore, when $e_1 \overset{R}{\leftarrow} [E_1]_n$,
\[ \Pr[A(e_1) = 0] = \]
\[ \Pr[p_{\text{psp}}(e_1, tT) \equiv p(E_0, \text{rec}(E_0))] = \]
\[ (1 - \Pr[p_{\text{psp}}(e_1, tT) \neq p(E_0, \text{rec}(E_0))]) \leq \]
\[ 1 - \Pr[p_{\text{psp}}(e_1, tT) \equiv p(E_1, \text{rec}(E_1))] = 1 - \Pr[A_1] \]

and we obtain:
\[ \Pr[A(e_1) = 0] \leq 2|E_1|^3 \nu_0(\eta) \]

(4)

From (3) and (4) if follows that $A$ distinguishes between $[E_0]_n$ and $[E_1]_n$ with probability greater than:
\[ 1 - 2(|E_0|^3 + |E_1|^3) \nu_0(\eta) \]

which is non-negligible.

4 Encryption schemes that leak the length of the message

In this section we show the validity of Theorems 2.1 and 3.3 for weaker but more reasonable security assumptions. More precisely we show that for encryption schemes which do not hide the length of the plaintext, a simple modification of the pattern semantics preserves the above mentioned results.

We make the simplifying assumption that basic messages are encoded in blocks of length equal to the length of the key. Also, we assume that a ciphertext is one block longer than the underlying plaintext. It is important to observe that these restrictions are not significant, since widely used symmetric encryption schemes such as CBC ([4]) have this property.

While the language of expressions remains the same, we modify the language of patterns by replacing $\square$ with natural numbers. The intuition is that the knowledge an adversary may have contains information of the type: “undeckryptable ciphertext of length $n$ blocks” rather than simply “undeckryptable ciphertext”. The grammar for patterns becomes:

\[ P := K \quad \text{key (for } K \in \text{Keys}) \]
\[ i \quad \text{bit (for } i \in \text{Bool}) \]
\[ (M, N) \quad \text{pair} \]
\[ \{M\}_K \quad \text{encryption (for } K \in \text{Keys}) \]
\[ n \quad \text{undeckryptable ciphertext of length } n \]

(5)

Next, consider the function $L : \text{Exp} \rightarrow \mathbb{N}$ defined inductively as follows:

- $L(K) = 1$ for any $K \in \text{Keys}$;
- $L(B) = 1$ for any $B \in \text{Blocks}$;
- $L((E_0, E_1)) = L(E_0) + L(E_1)$
  
  for any $E_0, E_1 \in \text{Exp}$;
- $L(\{E\}_K) = L(E) + 1$;

We change the definition of $p$ by replacing $p(\{E\}_K, T) = \square$ with $p(\{E\}_K, T) = L(E)$ for the case when $K \notin T$. Expression equivalence remains the same: two expressions are equivalent if their associated patterns are equal up to renaming. The computational semantics remains essentially the same as in [2].

It is not hard to see that while the statement and the proof of Theorem 3.3 remain essentially the same, the hypothesis of Theorem 2.1 can now be weaken to include “type-1” secure encryption schemes (encryption schemes hiding all partial information about
the keys and all partial information about encrypted plaintexts except possibly their length):

**Theorem 4.1** Let $E_0$ and $E_1$ be acyclic expressions and let $\Pi$ be a type-1 secure encryption scheme. Suppose that $E_0 \cong E_1$, then $[E_0]_{\Pi(i)} \cong [E_1]_{\Pi(i)}$.

**References**


