Equivalence of perfect privacy definitions

TEK4500 – Fall 2023
Håkon Jacobsen

There are a number of different formulations of one-time perfect privacy that all turn out to be equivalent to the game-based definition we gave in class. In this note we prove this equivalence for one of the most common formulations (also used in [BR]). First recall the definition of perfect privacy we used in class. For the purposes of this note, we call this definition *game-based* one-time perfect privacy.

Definition 1. An encryption scheme $\Sigma = (\mathsf{Enc}, \mathsf{Dec})$ defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ has (*game-based*) *one-time perfect privacy* if any adversary $\mathcal{A}$ has zero advantage in the following game

$\underline{\mathbf{Exp}^{\text{1-priv}}_{\Sigma}(\mathcal{A})}$

1: $b \xleftarrow{\$} \{0,1\}$
2: $K \xleftarrow{\$} \Sigma.\mathsf{KeyGen}$
3: $M \leftarrow \mathcal{A}$ // $\mathcal{A}$ picks message it wants to see encrypted
4: $R \xleftarrow{\$} \{0,1\}^{|M|}$ // Challenger draws random msg. of equal length
5: $C_0 \leftarrow \Sigma.\mathsf{Enc}(K, M)$ // "World 0": encrypt M
6: $C_1 \leftarrow \Sigma.\mathsf{Enc}(K, R)$ // "World 1": encrypt R
7: $b' \leftarrow \mathcal{A}(C_b)$ // $\mathcal{A}$ tries to guess which world it's in based on C_b
8: **return** $b' \stackrel{?}{=} b$

where the *advantage* of $\mathcal{A}$ is defined as

$$\mathbf{Adv}^{\text{1-priv}}_{\Sigma}(\mathcal{A}) \stackrel{\text{def}}{=} \left| \Pr[\mathbf{Exp}^{\text{1-priv}}_{\Sigma}(\mathcal{A}) \Rightarrow \mathsf{true}] - \frac{1}{2} \right|.$$

Here is the (more standard) definition of perfect privacy used in [BR].

Definition 2. An encryption scheme $\Sigma = (\mathsf{Enc}, \mathsf{Dec})$ defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ has *one-time perfect privacy* if for every $M, M' \in \mathcal{M}$ and every $C \in \mathcal{C}$

$$\Pr[\mathsf{Enc}(K, M) = C] = \Pr[\mathsf{Enc}(K, M') = C]. \tag{1}$$

In both cases the probability is taken over the random choice $K \xleftarrow{\$} \mathcal{K}$ and the coins tossed by Enc (if any).

Theorem 1. *Definition 1 and Definition 2 are equivalent definitions of one-time perfect privacy.*

Theorem 1 says it that if an encryption scheme is secure according to the game-based definition in Definition 1 then it is also secure according to the definition in Definition 2, and vice versa. We start by proving the direction Definition 2 $\implies$ Definition 1, that is, the (standard) one-time perfect privacy definition implies the game-based one.

Proof (*Definition 2* $\implies$ *Definition 1*)*:* This is basically just a generalization of the argument given in class: since $\Pr[\mathsf{Enc}_K(M) = C] = \Pr[\mathsf{Enc}_K(M') = C]$ the ciphertext that $\mathcal{A}$ sees is distributed identically in both World 1 and World 0, so it is impossible to distinguish the two. Hence $\Pr[b' = b] = 1/2$.

More formally, since we're dealing with unbounded adversaries we can without loss of generality assume that $\mathcal{A}$ is deterministic (why?). Now let $\mathcal{C}_0$ (resp. $\mathcal{C}_1$) denote the set of all the ciphertexts for which $\mathcal{A}$ outputs $b' = 0$ (resp. $b' = 1$). Note that since $\mathcal{A}$ must output either 0 or 1, we have $\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_1$ where $\mathcal{C}_0$ and $\mathcal{C}_1$ are mutually exclusive. Let $\Pr[\mathsf{Enc}_K(\$) = C]$ denote the probability that $\mathsf{Enc}_K(\widetilde{M})$ equals C for a randomly drawn message $\widetilde{M} \in \mathcal{M}$. Then:

$$
\begin{aligned}
\Pr[b' = b] &= \frac{1}{2} \cdot \Pr[b' = 1 \mid b = 1] + \frac{1}{2} \cdot \Pr[b' = 0 \mid b = 0] \\
&= \frac{1}{2} \sum_{C \in \mathcal{C}_1} \Pr[\mathsf{Enc}_K(M) = C] + \frac{1}{2} \sum_{C \in \mathcal{C}_0} \Pr[\mathsf{Enc}_K(\$) = C] \\
&\overset{a)}{=} \frac{1}{2} \sum_{C \in \mathcal{C}_1} \Pr[\mathsf{Enc}_K(M) = C] + \frac{1}{2} \sum_{C \in \mathcal{C}_0} \Pr[\mathsf{Enc}_K(M) = C] \\
&= \frac{1}{2} \sum_{C \in \mathcal{C}} \Pr[\mathsf{Enc}_K(M) = C] \\
&\overset{b)}{=} \frac{1}{2},
\end{aligned}
$$

which proves that Σ has perfect privacy according to the game-based definition (Definition 1). In $a)$ we used the assumption that Σ has perfect privacy according to Definition 2, hence $\Pr[\mathsf{Enc}_K(M) = C] = \Pr[\mathsf{Enc}_K(\$) = C]$. In $b)$ we used that $\sum_{C \in \mathcal{C}} \Pr[\mathsf{Enc}_K(M) = C] = 1$. □

We now prove the other direction: security according to the game-based definition (Definition 1) implies security according to the standard definition (Definition 2). However, when proving it's easier to instead prove the (logically equivalent) contrapositive statement: if Σ does *not* have perfect privacy according to Definition 2, then it also does *not* have perfect privacy according to Definition 1. In particular, if Σ does *not* have security according to Definition 2 then there must be *some* M, M', C for which $\Pr[\mathsf{Enc}_K(M) = C] \neq \Pr[\mathsf{Enc}_K(M') = C]$. From this fact we show that we can create an adversary that is able to

distinguish World 1 and World 0 in the game-based experiment with probability different from $1/2$, i.e., $\Pr[b' = b] \neq 1/2$. This is exactly what's needed to show that Σ is *not* secure according to Definition 1.

Proof (*Definition 1* $\implies$ *Definition 2*): Assume that Σ is *not* secure according to Definition 2. By definition, this means that there exist $M, M' \in \mathcal{M}$, $C \in \mathcal{C}$ such that $\Pr[\mathsf{Enc}_K(M) = C] \neq \Pr[\mathsf{Enc}_K(M') = C]$. Consequently, for this C there must also be an M for which $\Pr[\mathsf{Enc}_K(M) = C]$ is *maximal*[1], hence $\Pr[\mathsf{Enc}_K(M) = C] > \Pr[\mathsf{Enc}_K(M') = C]$. Note here that the probability is taken over the choice of all keys (and whatever random coins flipped by the encryption algorithm Enc), *not* over the choice of messages.

From the above fact we can construct the following game-based adversary $\mathcal{A}$ that wins in experiment $\mathbf{Exp}^{\text{1-priv}}_{\Sigma}(\mathcal{A})$ with probability strictly better than $1/2$ (or equivalently: with advantage $\mathbf{Adv}^{\text{1-priv}}_{\Sigma}(\mathcal{A}) > 0$). Adversary $\mathcal{A}$ simply submits the message M having maximal probability $\Pr[\mathsf{Enc}_K(M) = C]$ to the challenger and outputs $b' = 0$ if the returned ciphertext equals C, and outputs a random bit b' if not.[2]

What's the probability that $\mathcal{A}$ wins in $\mathbf{Exp}^{\text{1-priv}}_{\Sigma}(\mathcal{A})$? To simplify notation, let Win denote the event that $b' = b$ and let $\Pr[\mathsf{Enc}_K(\$) = C]$ denote the probability that $\mathsf{Enc}_K(\widetilde{M})$ equals C for a randomly drawn message $\widetilde{M} \in \mathcal{M}$. In the expression $\Pr[\mathsf{Enc}_K(\$) = C]$ the probability is taken over the choice of keys *and* the choice of messages (and whatever randomness used by Enc)—unlike in $\Pr[\mathsf{Enc}_K(M) = C]$ where the probability is only taken over the choice of keys (and whatever randomness used by Enc) not the choice of message (which is fixed).

In the real world we then have:

$$
\begin{aligned}
\Pr[\mathsf{Win} \mid b = 0] &= \Pr[\mathsf{Win} \mid b = 0 \wedge \mathsf{Enc}_K(M) = C] \cdot \Pr[\mathsf{Enc}_K(M) = C] \\
&\quad + \Pr[\mathsf{Win} \mid b = 0 \wedge \mathsf{Enc}_K(M) \neq C] \cdot \Pr[\mathsf{Enc}_K(M) \neq C] \\
&= 1 \cdot \Pr[\mathsf{Enc}_K(M) = C] + \frac{1}{2} \cdot \Pr[\mathsf{Enc}_K(M) \neq C] \\
&= \Pr[\mathsf{Enc}_K(M) = C] + \frac{1}{2} \cdot (1 - \Pr[\mathsf{Enc}_K(M) = C]) \\
&= \frac{1}{2} + \frac{1}{2} \cdot \Pr[\mathsf{Enc}_K(M) = C]
\end{aligned}
$$

[1] Note that there could be multiple messages attaining the maximum.

[2] How does $\mathcal{A}$ actually find M and C? This is where we use the fact that the perfect privacy definition allows $\mathcal{A}$ to be computationally *unbounded*: $\mathcal{A}$ can simply enumerate over *all* possible keys, messages, and ciphertexts, until it finds an M that maximizes $\Pr[\mathsf{Enc}_K(M) = C]$.

Similarly, in the ideal world, we have:

$$\begin{aligned}\Pr[\mathsf{Win} \mid b=1] &= \Pr[\mathsf{Win} \mid b=1 \wedge \mathsf{Enc}_K(\$) = C] \cdot \Pr[\mathsf{Enc}_K(\$) = C] \\ &\quad + \Pr[\mathsf{Win} \mid b=1 \wedge \mathsf{Enc}_K(\$) \neq C] \cdot \Pr[\mathsf{Enc}_K(\$) \neq C] \\ &= 0 \cdot \Pr[\mathsf{Enc}_K(\$) = C] + \frac{1}{2} \cdot \Pr[\mathsf{Enc}_K(\$) \neq C] \\ &= \frac{1}{2} \cdot (1 - \Pr[\mathsf{Enc}_K(\$) = C]) \\ &= \frac{1}{2} - \frac{1}{2} \cdot \Pr[\mathsf{Enc}_K(\$) = C]\end{aligned}$$

Putting these two together we get:

$$\begin{aligned}\Pr[\mathsf{Win}] &= \frac{1}{2} \cdot \Pr[\mathsf{Win} \mid b=0] + \frac{1}{2} \cdot \Pr[\mathsf{Win} \mid b=1] \\ &= \frac{1}{4} + \frac{1}{4} \cdot \Pr[\mathsf{Enc}_K(M) = C] + \frac{1}{4} - \frac{1}{4} \cdot \Pr[\mathsf{Enc}_K(\$) = C] \\ &= \frac{1}{2} + \frac{1}{4} \cdot \Pr[\mathsf{Enc}_K(M) = C] - \frac{1}{4} \cdot \sum_{\widetilde{M} \in \mathcal{M}} \Pr[\mathsf{Enc}_K(\widetilde{M}) = C] \cdot \frac{1}{|\mathcal{M}|} \\ &= \frac{1}{2} + \frac{1}{4 \cdot |\mathcal{M}|} \cdot \Big(|\mathcal{M}| \cdot \Pr[\mathsf{Enc}_K(M) = C] \\ &\qquad - \sum_{\substack{\widetilde{M} \in \mathcal{M} \\ \widetilde{M} \neq M'}} \Big(\Pr[\mathsf{Enc}_K(\widetilde{M}) = C] \Big) - \Pr[\mathsf{Enc}_K(M') = C] \Big) \\ &\overset{a)}{\geq} \frac{1}{2} + \frac{1}{4 \cdot |\mathcal{M}|} \cdot \Big(|\mathcal{M}| \cdot \Pr[\mathsf{Enc}_K(M) = C] \\ &\qquad - (|\mathcal{M}| - 1) \cdot \Pr[\mathsf{Enc}_K(M) = C] - \Pr[\mathsf{Enc}_K(\widetilde{M}) = C] \Big) \\ &= \frac{1}{2} + \frac{1}{4 \cdot |\mathcal{M}|} \cdot \big(\Pr[\mathsf{Enc}_K(M) = C] - \Pr[\mathsf{Enc}_K(M') = C] \big) \\ &\overset{b)}{>} \frac{1}{2}. \end{aligned} \tag{2}$$

In $a)$ we used that $\Pr[\mathsf{Enc}_K(M) = C]$ is maximal, hence

$$(|\mathcal{M}| - 1) \cdot \Pr[\mathsf{Enc}_K(M) = C] \geq \sum_{\substack{\widetilde{M} \in \mathcal{M} \\ \widetilde{M} \neq M'}} \Pr[\mathsf{Enc}_K(\widetilde{M}) = C].$$

In $b)$ we used the assumption that $\Pr[\mathsf{Enc}_K(M) = C] > \Pr[\mathsf{Enc}_K(M') = C]$, thus

$$\frac{1}{4 \cdot |\mathcal{M}|} \cdot \big(\Pr[\mathsf{Enc}_K(M) = C] - \Pr[\mathsf{Enc}_K(M') = C] \big) > 0.$$

Finally, this means that the *advantage* of $\mathcal{A}$ is > 0, since

$$\mathbf{Adv}_{\Sigma}^{\text{1-priv}}(\mathcal{A}) \stackrel{\text{def}}{=} \left| \Pr[\mathsf{Win}] - \frac{1}{2} \right| > 0,$$

by (2). This proves that Σ does not have perfect privacy according to Definition 1, showing that $\neg$ Definition 2 $\implies$ $\neg$ Definition 1. By the contrapositive, this means that Definition 1 $\implies$ Definition 2, which is what we wanted to prove. □

References

[BR] Mihir Bellare and Phillip Rogaway. *Introduction to Modern Cryptography*. https://web.cs.ucdavis.edu/~rogaway/classes/227/spring05/book/main.pdf.