ETH Zürich, D-INFK
Spring 2018

Prof. Ueli Maurer
Christian Badertscher
Fabio Banfi

Cryptography Foundations
Solution Exercise 11

11.1 Conditional Probability Distributions

a) We have

$$\mathsf{p}^{\mathbf{E}^\perp}_{X_1}(x_1) = \frac{1}{|\{0,1\}^n|} = 2^{-n},$$
$$\mathsf{p}^{\mathbf{E}^\perp}_{X_2|X_1Y_1}(x_2, x_1, y_1) = \begin{cases} 0, & x_2 \neq y_1 \\ 1, & x_2 = y_1, \end{cases}$$

and for $i > 2$, $y^i = (y_1, \ldots, y_i)$, and $x^{i-1} = (x_1, \ldots, x_{i-1})$ with $x_2 = y_1$ and $x_k = \perp$ for all $3 \leq k \leq i-1$,

$$\mathsf{p}^{\mathbf{E}^\perp}_{X_i|X^{i-1}Y^{i-1}}(x_i, x^{i-1}, y^{i-1}) = \begin{cases} 0, & x_i \neq \perp \\ 1, & x_i = \perp. \end{cases}$$

It follows that

$$\mathsf{p}^{\mathbf{E}^\perp}_{X_1X_2|Y_1}(x_1, x_2, y_1) = \mathsf{p}^{\mathbf{E}^\perp}_{X_1}(x_1) \cdot \mathsf{p}^{\mathbf{E}^\perp}_{X_2|X_1Y_1}(x_2, x_1, y_1) = \begin{cases} 0, & x_2 \neq y_1 \\ 2^{-n}, & x_2 = y_1. \end{cases}$$

This yields using the behavior of a URP:

$$\begin{aligned} \mathsf{P}^{\mathbf{EP}_n}_{X_1X_2Y_1Y_2}(x_1, x_2, y_1, y_2) &= \mathsf{p}^{\mathbf{E}^\perp}_{X_1X_2|Y_1}(x_1, x_2, y_1) \cdot \mathsf{p}^{\mathbf{P}_n}_{Y_1Y_2|X_1X_2}(y_1, y_2, x_1, x_2) \\ &= \begin{cases} 0, & x_2 \neq y_1 \vee (x_2 = x_1 \wedge y_2 \neq y_1) \vee (x_2 \neq x_1 \wedge y_2 = y_1) \\ \frac{1}{2^{2n}}, & x_1 = x_2 = y_1 = y_2 \\ \frac{1}{2^{2n}} \cdot \frac{1}{2^n - 1}, & \text{else} \end{cases} \end{aligned}$$

b) We can compute the marginal distribution as follows:

$$\begin{aligned} \mathsf{P}^{\mathbf{EP}_n}_{Y_1Y_2}(y_1, y_2) &= \sum_{x_1 \in \{0,1\}^n} \sum_{x_2 \in \{0,1\}^n} \mathsf{P}^{\mathbf{EP}_n}_{X_1X_2Y_1Y_2}(x_1, x_2, y_1, y_2) \\ &= \sum_{x_1 \in \{0,1\}^n} \mathsf{P}^{\mathbf{EP}_n}_{X_1X_2Y_1Y_2}(x_1, y_1, y_1, y_2) \\ &= \begin{cases} \frac{1}{2^{2n}}, & y_1 = y_2, \\ (2^n - 1) \cdot \frac{1}{2^{2n}} \cdot \frac{1}{2^n - 1}, & y_1 \neq y_2 \end{cases} \\ &= 2^{-2n} \end{aligned}$$

Since

$$\mathsf{P}^{\mathbf{EP}_n}_{X_1X_2}(x_1, x_2) = \mathsf{P}^{\mathbf{EP}_n}_{X_1Y_1}(x_1, x_2) = \mathsf{p}^{\mathbf{E}^\perp}_{X_1}(x_1) \cdot \mathsf{p}^{\mathbf{P}_n}_{Y_1|X_1}(y_1, x_1) = 2^{-n} \cdot 2^{-n} = 2^{-2n},$$

we finally get:

$$\mathsf{P}^{\mathbf{EP}_n}_{Y_1Y_2|X_1X_2}(y_1,y_2,x_1,x_2) = \frac{\mathsf{P}^{\mathbf{EP}_n}_{X_1X_2Y_1Y_2}(x_1,x_2,y_1,y_2)}{\mathsf{P}^{\mathbf{EP}_n}_{X_1X_2}(x_1,x_2)}$$
$$= \begin{cases} 0, & x_2 \neq y_1 \vee (x_2 = x_1 \wedge y_2 \neq y_1) \vee (x_2 \neq x_1 \wedge y_2 = y_1) \\ 1, & x_1 = x_2 = y_1 = y_2 \\ \frac{1}{2^n-1}, & \text{else.} \end{cases}$$

Note that we have $\mathsf{P}^{\mathbf{EP}_n}_{Y_1Y_2|X_1X_2} \neq \mathsf{p}^{\mathbf{P}_n}_{Y_1Y_2|X_1X_2}$. In general, the conditional probabilities in a random experiment are not necessarily equal to the values obtained by the "small p's" describing the behavior of the systems. However, one can show that if the environment chooses the inputs X_i to the system independently of the outputs Y_i, the conditional probabilities are equal.

11.2 Distinguishing URFs and URPs

a) As suggested in Example 4.19, we can define the MBO that remains 0 as long as for two distinct inputs, the outputs are distinct. More formally:

$$A_i := \begin{cases} 0, & \forall\ 1 \leq j,j' \leq i \quad x_j = x_{j'} \ \vee \ y_j \neq y_{j'} \\ 1, & \exists\ 1 \leq j,j' \leq i \quad x_j \neq x_{j'} \ \wedge \ y_j = y_{j'} \end{cases}$$

To proof $\hat{\mathbf{R}}_{n.n} \models \mathbf{P}_n$, one only has to consider distinct inputs, since both systems answer repeating queries consistently. Conditioned on $A_i = 0$, $\hat{\mathbf{R}}_{n,n}$ produces a uniformly random vector y^i with distinct entries for any input vector x^i. Since the same holds for $\mathbf{P}_n$, we intuitively have $\hat{\mathbf{R}}_{n.n} \models \mathbf{P}_n$.
We now provide a formal proof based on this intuition: For a sequence of i inputs x^i, let $\mathsf{d}(x^i)$ be the number of distinct elements. We get

$$\mathsf{p}^{\hat{\mathbf{R}}_{n,n}}_{Y^i,A_i=0|X^i}(y^i,x^i) = \prod_{j=1}^{i} \mathsf{p}^{\hat{\mathbf{R}}_{n,n}}_{Y_j,A_j=0|X^jY^{j-1},A_{j-1}=0}(y_j,x^j,y^{j-1})$$
$$= \begin{cases} 0, & \exists\ 1 \leq j,j' \leq i \quad (x_j = x_{j'} \wedge y_j \neq y_{j'}) \vee (x_j \neq x_{j'} \ \wedge \ y_j = y_{j'}) \\ 2^{-n\mathsf{d}(x^i)}, & \text{else.} \end{cases}$$

Note that $\mathsf{p}^{\hat{\mathbf{R}}_{n,n}}_{A_i=0|X^i}(x^i)$ is the probability that for all distinct x_i's, the corresponding y_i's are distinct. Hence,

$$\mathsf{p}^{\hat{\mathbf{R}}_{n,n}}_{A_i=0|X^i}(x^i) = \prod_{j=0}^{\mathsf{d}(x^i)-1} \frac{2^n-j}{2^n} \ = \ 2^{-n\mathsf{d}(x^i)} \cdot \prod_{j=0}^{\mathsf{d}(x^i)-1} (2^n-j).$$

Recall from the 6th exercise sheet that

$$\mathsf{p}^{\mathbf{P}_n}_{Y^i|X^i}\left(y^i,x^i\right) = \begin{cases} 0, & \exists\ 1 \leq j,j' \leq i \quad (x_j = x_{j'} \wedge y_j \neq y_{j'}) \\ & \qquad \vee\ (x_j \neq x_{j'} \wedge y_j = y_{j'}), \\ \prod_{j=0}^{\mathsf{d}(x^i)-1} \frac{1}{2^n-j}, & \text{else.} \end{cases}$$

We therefore have

$$\mathsf{p}^{\hat{\mathbf{R}}_{n,n}}_{Y^i,A_i=0|X^i}(y^i,x^i) = \mathsf{p}^{\hat{\mathbf{R}}_{n,n}}_{A_i=0|X^i}(x^i) \cdot \mathsf{p}^{\mathbf{P}_n}_{Y^i|X^i}\left(y^i,x^i\right),$$

which is equivalent to $\hat{\mathbf{R}}_{n.n} \models \mathbf{P}_n$.

b) We have $\Pr^{X_i X_j}[X_i = X_j] = \frac{1}{t}$ for $i \neq j$ and there are $\binom{q}{2} = \frac{q(q-1)}{2}$ such pairs. We conclude using the union bound

$$\begin{aligned}
\mathsf{p}_{\text{coll}}(q,t) &= \Pr^{X_1 \dots X_q}[\exists i,j \quad 1 \leq i < j \leq q \ \wedge \ X_i = X_j] \\
&= \Pr^{X_1 \dots X_q}\left[\bigcup_{1 \leq i < j \leq q} X_i = X_j\right] \\
&\leq \sum_{1 \leq i < j \leq q} \Pr^{X_i X_j}[X_i = X_j] \\
&= \frac{q(q-1)}{2t} \\
&\leq \frac{1}{2} q^2/t.
\end{aligned}$$

11.3 Distinguishing Systems Adaptively

a) The answer to the first input is uniformly distributed for both systems. Hence, there is no way to tell the two systems apart since $[1]\mathbf{S}_0 \equiv [1]\mathbf{S}_1$.

b) We first input some value $x_1 \in \{0,1\}^n$. Given the corresponding output $y_1 \in \{0,1\}^n$, we input $x_2 = y_1$ as our second input. System $\mathbf{S}_0$ always returns some value $y_2 \in \{0,1\}^n$, while $\mathbf{S}_1$ always returns $\perp$ in this case. Therefore, we are *always* able to distinguish.

c) If we have to fix our inputs $x_1, \dots, x_k$ before we get to see the corresponding outputs $y_1, \dots, y_k$, we only notice a difference if, for some i, we have $y_i \in \{x_{i+1}, \dots, x_k\}$. In this case, we get to see an output $\perp$ in the case where we are interacting with $\mathbf{S}_1$. For some fixed i, the probability that this happens is at most $(k-i) \cdot 2^{-n} < k \cdot 2^{-n}$, and the probability that this happens for *some* i is at most $k \cdot k \cdot 2^{-n} = k^2 \cdot 2^{-n}$. Here, we used the union bound twice.

That is, if k is small ($k < 2^{n/2}$) then guessing the correct system is hard, as $\mathbf{S}_1$ behaves like $\mathbf{S}_0$ with high probability.

We can make the above argument using the concepts from the lecture and model the above situation as follows: we consider the induced systems $\mathbf{S}'_\ell$ (with $\ell \in \{0,1\}$) which take as input exactly one vector $(x_1, \dots, x_k) \in (\{0,1\}^n)^k$, interact (as an environment) with the system $\mathbf{S}_\ell$ to produce the output vector $(y_1, \dots, y_k)$. Define the binary event E_i in this interaction such that E_i occurs if and only if $y_i \in \{x_{i+1}, \dots, x_k\}$. Let A be the MBO of the induced systems $\hat{\mathbf{S}}'_\ell$ such that $A_1 = 1$ if at least one of the events E_i occurred (and $A_t := A_{t-1}$ for $t > 1$). We have that $\hat{\mathbf{S}}'_0 \overset{g}{\equiv} \hat{\mathbf{S}}'_1$ since $\mathsf{p}^{\hat{\mathbf{S}}'_0}_{Y,A_1=0|X}((y^k),(x^k)) = \mathsf{p}^{\hat{\mathbf{S}}'_1}_{Y,A_1=0|X}((y^k),(x^k))$ (which can be verified as an additional exercise similar to 11.1). Hence, Lemma 16 implies that the advantage of the best distinguisher for $\mathbf{S}'_0$ and $\mathbf{S}'_1$ is upper bounded by the probability of provoking $A_1 = 1$ (of system $\hat{\mathbf{S}}'_1$) which is no more than $k^2 \cdot 2^{-n}$ by the union-bound argument from above.

11.4 Expansion of PRGs

a) We aim at constructing the specification

$$\mathcal{S} := \{\mathbf{U}_{4k}\}^f = \{\mathbf{S} \in \Phi \mid \overline{\langle \mathbf{S} \mid \mathbf{U}_{4k} \rangle} \leq f\},$$

where $f = \lambda\, \overline{\langle \alpha[\mathbf{U}_k, \mathbf{G}] \mid \mathbf{U}_{2k} \rangle}\, \rho$.

b) For this task, assume that $\mathbf{U}_\kappa$, for any $\kappa \in \mathbb{N}$, accepts inputs from $\{\diamond\} \cup \{0,1\}^*$, and any such input is treated as the trigger input $\diamond$, therefore $\mathbf{U}_\kappa$ essentially ignores its input and

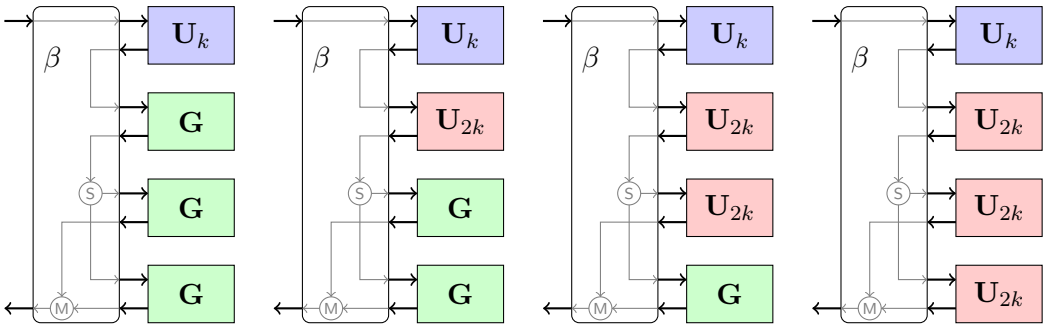

Figure 1: Description of the converter β and outline of the hybrid argument. The node S splits a $2k$-bit string into two k-bit strings, whereas the node M merges two k-bit strings into a $2k$-bit string. Note that since $\mathbf{U}_k$ and $\mathbf{U}_{2k}$ accept any input as trigger, they essentially ignore their inputs.

replies with a fresh uniform κ-bit string. The assumed specification is the singleton set $\mathcal{R} = \{[\mathbf{U}_k, \mathbf{G}, \mathbf{G}, \mathbf{G}]\}$. The converter β is described in Figure 1.

We now introduce the following hybrids, parameterized on a resource $\mathbf{X}$:

$$\begin{aligned}\mathbf{H}_1(\mathbf{X}) &:= \beta[\mathbf{U}_k, \mathbf{X}, \mathbf{G}, \mathbf{G}],\\ \mathbf{H}_2(\mathbf{X}) &:= \beta[\mathbf{U}_k, \mathbf{U}_{2k}, \mathbf{X}, \mathbf{G}],\\ \mathbf{H}_3(\mathbf{X}) &:= \beta[\mathbf{U}_k, \mathbf{U}_{2k}, \mathbf{U}_{2k}, \mathbf{X}].\end{aligned}$$

Note that for any $i \in \{1, 2\}$ and $j \in \{1, 2, 3\}$, we have

$$\mathbf{H}_i(\mathbf{U}_{2k}) \equiv \mathbf{H}_{i+1}(\mathbf{G}) \qquad \text{and} \qquad \mathbf{H}_j(\mathbf{G}) \equiv \mathbf{H}_j(\alpha[\mathbf{U}_k, \mathbf{G}]), \tag{1}$$

where the second equivalence follows from the fact that the input to the resource $\mathbf{G}$ given as argument to $\mathbf{H}_j$ is a k-bit uniform random string, and therefore this source can be ignored and $\mathbf{G}$ replaced by $\alpha[\mathbf{U}_k, \mathbf{G}]$ (which simply corresponds to "moving" the randomness source fed to $\mathbf{G}$). Also note that we the ideal resource $\mathbf{U}_{4k}$ is equivalent to $\beta[\mathbf{U}_k, \mathbf{U}_{2k}, \mathbf{U}_{2k}, \mathbf{U}_{2k}]$, and therefore, using the hybrids, we get

$$\overline{\langle \beta[\mathbf{U}_k, \mathbf{G}, \mathbf{G}, \mathbf{G}] \,|\, \mathbf{U}_{4k} \rangle} = \overline{\langle \mathbf{H}_1(\mathbf{G}) \,|\, \mathbf{H}_3(\mathbf{U}_{2k}) \rangle}.$$

Using Lemma 2.2 and Equation 1, we obtain

$$\overline{\langle \mathbf{H}_1(\mathbf{G}) \,|\, \mathbf{H}_3(\mathbf{U}_{2k}) \rangle} = \sum_{i=1}^{3} \overline{\langle \mathbf{H}_i(\mathbf{G}) \,|\, \mathbf{H}_i(\mathbf{U}_{2k}) \rangle} = \sum_{i=1}^{3} \overline{\langle \mathbf{H}_i(\alpha[\mathbf{U}_k, \mathbf{G}]) \,|\, \mathbf{H}_i(\mathbf{U}_{2k}) \rangle}.$$

Finally, with I uniformly distributed over $\{1, 2, 3\}$, we have

$$\begin{aligned}\sum_{i=1}^{3} \overline{\langle \mathbf{H}_i(\alpha[\mathbf{U}_k, \mathbf{G}]) \,|\, \mathbf{H}_i(\mathbf{U}_{2k}) \rangle} &= \sum_{i=1}^{3} \overline{\langle \alpha[\mathbf{U}_k, \mathbf{G}] \,|\, \mathbf{U}_{2k} \rangle} \rho^{\mathbf{H}_i(\cdot)}\\ &= 3 \cdot \sum_{i=1}^{3} \Pr[I = i] \cdot \overline{\langle \alpha[\mathbf{U}_k, \mathbf{G}] \,|\, \mathbf{U}_{2k} \rangle} \rho^{\mathbf{H}_I(\cdot)}\\ &= 3 \cdot \overline{\langle \alpha[\mathbf{U}_k, \mathbf{G}] \,|\, \mathbf{U}_{2k} \rangle} \rho^{\mathbf{H}_I(\cdot)},\end{aligned}$$

where we used the law of total probability and the fact that $\Pr[I = i] = \frac{1}{3}$.

Therefore, with $\lambda = 3$ and the reduction $\rho = \rho^{\mathbf{H}_I}$ as described above, we have $\mathcal{R} \subseteq \mathcal{S}$.