The art of cryptography: Heads and tails – Cryptographic random generation summer 2015

Distinguishers

Prof. Dr. Joachim von zur Gathen
Dr. Daniel Loebenberger

Let $\mathcal{A}$ be the deterministic algorithm which outputs $\mathcal{A}((x_1, \ldots, x_6)) = x_3$ on any input $(x_1, \ldots, x_6) \in \mathbb{B}^6$. Then for the uniform random variable U_6 on $\mathbb{B}^6$ we have

$$\mathcal{E}_{\mathcal{A}}(U_6) = \mathrm{prob}\{1 \xleftarrow{\text{🎲}} \mathcal{A}(U_6)\} = \mathrm{prob}\{1 \xleftarrow{\text{🎲}} U_1\} = \frac{1}{2}.$$

The U_1 here is the third component of $U_6 = U_1 \times U_1 \times U_1 \times U_1 \times U_1 \times U_1 = U_1^6$.

DEFINITION. *We take two distributions X_0 and X_1 on the same finite set A and an algorithm $\mathcal{A}$ which on input an element from A returns 0 or 1. Then the advantage $\sigma_{\mathcal{A}}(X_0, X_1)$ of $\mathcal{A}$ on X_1 over X_0 (or its distinguishing power between X_0 and X_1) is*

$$\sigma_{\mathcal{A}}(X_0, X_1) = |\mathcal{E}_{\mathcal{A}}(X_0) - \mathcal{E}_{\mathcal{A}}(X_1)|.$$

If $\sigma_{\mathcal{A}}(X_0, X_1) \geq \varepsilon > 0$, then $\mathcal{A}$ is an ε-distinguisher between X_0 and X_1, and X_0 and X_1 are ε-distinguishable.

Suppose that n is even and X takes only values with exactly $n/2$ ones: if $x \in \mathbb{B}^n$ and $\mathrm{prob}\{x \xleftarrow{\$} X\} > 0$, then $w(x) = \frac{n}{2}$. Here $w(x)$ is the Hamming weight of x, that is, the number of ones in x. Then the following deterministic algorithm $\mathcal{A}$ distinguishes between X and the uniform variable U_n on $\mathbb{B}^n$: on input x, return 1 if $w(x) = n/2$ else return 0. We have

$$\mathcal{E}_{\mathcal{A}}(X) = \mathrm{prob}\{1 \xleftarrow{\$} \mathcal{A}(X)\} = \mathrm{prob}\{\frac{n}{2} \xleftarrow{\$} w(X)\} = 1,$$

$$\begin{aligned}\mathcal{E}_{\mathcal{A}}(U_n) &= \mathrm{prob}\{1 \xleftarrow{\$} \mathcal{A}(U_n)\} = \mathrm{prob}\left\{\frac{n}{2} \xleftarrow{\$} w(U_n)\right\} \\ &= 2^{-n} \cdot \#\{x \in \mathbb{B}^n \colon w(x) = \frac{n}{2}\} = 2^{-n}\binom{n}{n/2}.\end{aligned}$$

Thus

$$|\mathcal{E}_{\mathcal{A}}(X) - \mathcal{E}_{\mathcal{A}}(U_n)| \approx 1 - \frac{1}{\sqrt{\pi n/2}} \geq \varepsilon$$

Definition. *A bit generator (or generator for short) is a function* $f : \mathbb{B}^k \longrightarrow \mathbb{B}^n$ *for some* $k < n$. *The corresponding random variable on* $\mathbb{B}^n$ *is* $f(U_k)$.

We consider the generator

$$f : \mathbb{B}^3 \longrightarrow \mathbb{B}^6$$

given by the following table

x	$f(x)$
000	001101
001	001011
010	011010
011	010110
100	101100
101	100101
110	110100
111	110010

We can easily distinguish the random variable $X = f(U_3)$ from U_6 by the distinguisher $\mathcal{A}$ from the previous example.

Define algorithm $\mathcal{B}$ as

$$\mathcal{B}(y) = \begin{cases} 1 & \text{if } y_4 = \text{minority}(y_1, y_2, y_3), \\ 0 & \text{otherwise.} \end{cases}$$

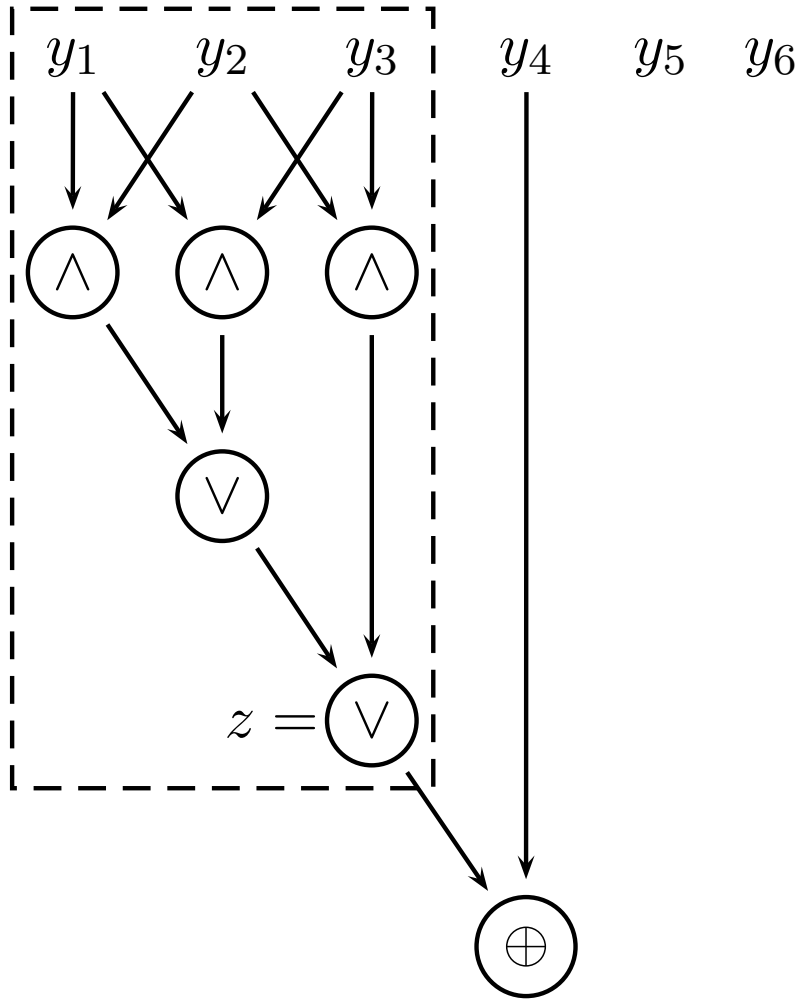
We now calculate $\mathcal{E}_{\mathcal{B}}(X) = \text{prob}\{1 \xleftarrow{\$} \mathcal{B}(X)\}$. There are eight values of y which occur as values of X, each with probability $1/8$.

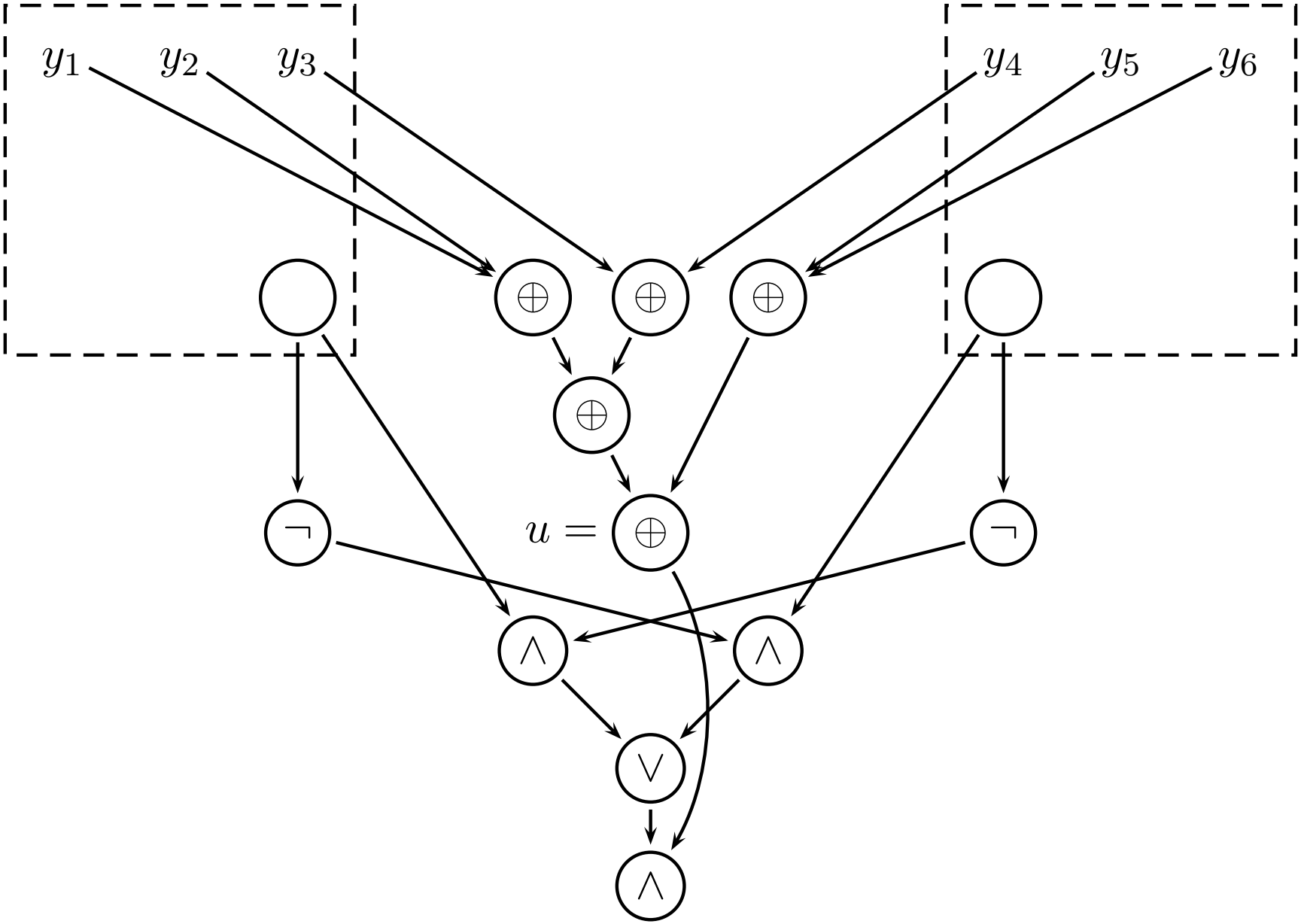
y	$\text{prob}\{1 \xleftarrow{\$} \mathcal{B}(y)\}$
001101	1
001011	0
011010	1
010110	1
101100	0
100101	1
110100	0
110010	1

DEFINITION. *A family* $g = (g_k)_{k\in\mathbb{N}}$ *with* $g_k \colon \mathbb{B}^k \longrightarrow \mathbb{B}^{n(k)}$ *and* $n(k) > k$ *for all* k *is a pseudorandom generator if*

- *it can be implemented in time polynomial in* k*,*
- *for all probabilistic polynomial-time algorithms* $\mathcal{A}$ *using polynomially many samples, the advantage* $\mathrm{Adv}_{\mathcal{A}}(U_{n(k)}, g_k(U_k))$ *of* $\mathcal{A}$ *on* $g_k(U_k)$ *is a negligible function of* k*.*

Definition. *Let $k < n$ and s be integers, $\varepsilon \geq 0$ real, and $f : \mathbb{B}^k \longrightarrow \mathbb{B}^n$ a generator. A probabilistic Boolean circuit $\mathcal{C}$ of size s and with advantage $\mathrm{Adv}_{\mathcal{C}}(U_n, f(U_k)) \geq \varepsilon$ is called an (ε, s)-distinguisher between U_n and $f(U_k)$. The circuit $\mathcal{C}$ has several inputs from $\mathbb{B}^n$, and either all of them are uniform or all drawn from $f(U_k)$. The function f is called an (ε, s)-resilient pseudorandom generator if no such $\mathcal{C}$ exists.*


y_1
y_2
y_3
y_4
y_5
y_6
$\wedge$
$\wedge$
$\wedge$
$\vee$
$z =$
$\vee$
$\oplus$


y_1
y_2
y_3
y_4
y_5
y_6
$u =$
$\oplus$
$\neg$
$\wedge$
$\vee$