MTAT.07.003 Cryptology II
Spring 2010 / Exercise session III / Example solution

Problem. Normally, it is impossible to compute computational distance between two distributions directly, since the number of potential distinguishing algorithms is humongous. However, for really small time-bounds it can be done. Assume that all distinguishers $\mathcal{A} : \mathbb{Z}_{16} \to \{0, 1\}$ are implemented as Boolean circuits consisting of Not, And, Or gates and the corresponding time-complexity is just the number of logic gates. For example, $\mathcal{A}(x_3x_2x_1x_0) = x_1$ has time-complexity 0 and $\mathcal{A}(x_3x_2x_1x_0) = x_1 \vee \neg x_3 \wedge x_2$ has time-complexity 3.

1. Let $\mathcal{X}_0$ be a uniform distribution over $\mathbb{Z}_{16}$ and let $\mathcal{X}_1$ be a uniform distribution over $\{0, 2, 4, 6, 8, 10, 12, 14\}$. What is $\mathsf{cd}_x^1(\mathcal{X}_0, \mathcal{X}_1)$?
2. Find a uniform distribution $\mathcal{X}_2$ over some 8 element set such that $\mathsf{cd}_x^1(\mathcal{X}_0, \mathcal{X}_2)$ is minimal. Compute $\mathsf{cd}_x^2(\mathcal{X}_0, \mathcal{X}_2)$ and $\mathsf{cd}_x^3(\mathcal{X}_0, \mathcal{X}_2)$.
3. Find a uniform distribution $\mathcal{X}_3$ over some 8 element set such that the distance sum $\mathsf{cd}_x^1(\mathcal{X}_1, \mathcal{X}_0) + \mathsf{cd}_x^1(\mathcal{X}_0, \mathcal{X}_3) \neq \mathsf{cd}_x^1(\mathcal{X}_1, \mathcal{X}_3)$.
4. Estimate for which value of t the distances $\mathsf{cd}_x^t(\mathcal{X}_0, \mathcal{X}_1)$ and $\mathsf{sd}_x(\mathcal{X}_0, \mathcal{X}_1)$ coincide for all distributions over $\mathbb{Z}_{16}$.

Solution. As the statistical distance $\mathsf{sd}_x(\mathcal{X}_0, \mathcal{X}_1) = \frac{1}{2}$ and the corresponding distinguisher $A(x_3x_2, x_1x_0) = x_0$ consists of zero gates, we get $\mathsf{cd}_x^0(\mathcal{X}_0, \mathcal{X}_1) = \frac{1}{2}$. For the second question, let $\mathcal{X}_\phi = \{x \in \mathbb{Z}_{16} : \phi(x) = 1\}$ denote the true-set for a circuit ϕ and let $\mathcal{X}_2$ be some 8 element set. Then by definition

$$
\begin{aligned}
\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{X}_0,\mathcal{X}_2}(\phi) &= |\Pr[x \leftarrow \mathcal{X}_0 : \phi(x) = 1] - \Pr[x \leftarrow \mathcal{X}_2 : \phi(x) = 1]| \\
&= \frac{1}{16} \cdot ||\mathcal{X}_\phi| - 2 \cdot |\mathcal{X}_\phi \cap \mathcal{X}_2|| = \frac{1}{16} \cdot ||\mathcal{X}_\phi| - |\mathcal{X}_\phi \setminus \mathcal{X}_2||
\end{aligned}
$$

and minimal computational distance is achieved by the set $\mathcal{X}_2$ that splits almost evenly by all possible sets $\mathcal{X}_\phi$. By considering formulae

$$\phi_1(x) = x_0, \ldots, \phi_4(x) = x_3, \phi_5(x) = \neg x_0, \ldots, \phi_8(x) = \neg x_3 \ ,$$

we get that a set $\mathcal{X}_2$ can achieve $\mathsf{cd}_x^1(\mathcal{X}_0, \mathcal{X}_2) = 0$ only if it contains 4 elements with the ith bit set to one and 4 elements with the ith bit set to zero. Formulae

$$
\begin{aligned}
\phi_9(x) = x_0 \wedge x_1,\ \ \phi_{10}(x) = x_0 \wedge x_2 \ldots, \phi_{13}(x) = x_1 \wedge x_3,\ \ \phi_{14}(x) = x_2 \wedge x_3 \ , \\
\phi_{15}(x) = x_0 \vee x_1,\ \ \phi_{16}(x) = x_0 \vee x_2 \ldots, \phi_{19}(x) = x_1 \vee x_3,\ \ \phi_{20}(x) = x_2 \vee x_3
\end{aligned}
$$

indicate that such a set must contain exactly 2 elements with ith and jth bit set to one and exactly 2 elements with ith and jth bit set to zero. A bit representation of a possible solution is depicted in Figure 1. The solution has a peculiar property: if we choose uniformly element from $\mathcal{X}_2$ and observe a bit

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	1	1	1	0	0	0	0
1	1	0	0	1	1	0	0
1	0	1	0	1	0	1	0
1	0	1	0	0	1	0	1
16	3	13	1	6	10	4	8

Figure 1: Orhogonal array with parameters $n = 4$ and $k = 2$.

pair i and j the corresponding bit-string has uniform distribution over $\mathbb{Z}_4$. Consequently, any formula consisting of two inputs is incapable from distinguishing $\mathcal{X}_0$ and $\mathcal{X}_2$. A formula consisting of two gates can cover three inputs and thus potential distinguishing capabilities are higher. As Figure 2 clearly shows, the distribution of bit triples x_0, x_2, x_3 is indeed different from uniform and the task of building a distinguisher simplifies considerably. In fact, we can express

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{X}_0,\mathcal{X}_2}(\phi) = \frac{1}{8} \cdot |\psi(000) + \psi(101) + \psi(110) - \psi(001) - \psi(100) - \psi(111)| \enspace .$$

for any formula $\phi(x) = \psi(x_0x_2x_3)$. Exhaustive search reveals that the formulae

$$x_0 \wedge x_2 \wedge x_3 \enspace , x_0 \vee x_2 \vee x_3 \enspace , x_0 \wedge x_3 \vee x_2, \enspace , x_0 \wedge (x_2 \vee x_3)$$

all achieve optimal advantage $\mathsf{Adv}^{\mathsf{ind}}_{\mathcal{X}_0,\mathcal{X}_2}(\phi) = \frac{1}{8}$. For the next distance estimate, note that a three gate distinguisher can cover all 4 inputs if it does not contain NOT-gates. All of such distinguishers achieve advantage $\frac{1}{16}$ and thus cannot not be optimal. Consequently, a potential optimal 3-gate distinguisher with NOT-gate must process inputs x_0, x_2, x_3. Indeed, several formulae with negation achieve again the advantage $\frac{1}{8}$ but not more. Hence, we have shown that

$$\mathsf{cd}^2_x(\mathcal{X}_0, \mathcal{X}_2) = \mathsf{cd}^3_x(\mathcal{X}_0, \mathcal{X}_2) = \frac{1}{8} \enspace .$$

As $\mathsf{sd}_x(\mathcal{X}_1, \mathcal{X}_1) = 0$ and $\mathsf{sd}_x(\mathcal{X}_0, \mathcal{X}_1) = \frac{1}{2}$, by taking $\mathcal{X}_3 = \mathcal{X}_1$ we get the required counter-example for the third question. Finally, note that any statistical test is a predicate. As a distinguisher with negated output works as well as the original, we must bound the gate complexity of a predicate that is satisfied by at most 8 inputs. Each of this inputs can be represented as conjunct consisting of three AND- and at most four NOT-gates. Hence, the total gate count is bounded by 64 gates, i.e., $\mathsf{cd}^{64}_x(\mathcal{X}_0, \mathcal{X}_1) = \mathsf{sd}_x(\mathcal{X}_0, \mathcal{X}_1)$ for all distributions $\mathcal{X}_0$ and $\mathcal{X}_1$.

Inputs	Violating triples	sd
x_0, x_1, x_2	No violating triples	0
x_0, x_1, x_3	No violating triples	0
x_0, x_2, x_3	$000 \to 0.00, 001 \to 0.25, 100 \to 0.25$	$\frac{3}{8}$
	$101 \to 0.00, 110 \to 0.00, 111 \to 0.25$	
x_2, x_3, x_4	No violating triples	0

Figure 2: Violating triples