MTAT.07.003 Cryptology II
Spring 2012 / Exercise session ?? / Example Solution

Exercise (Random self-reducibility of DL). *Let $\mathbb{G} = \langle g \rangle$ be a finite group of an order q generated by the powers of an element g. Then the Discrete Logarithm (DL) problem is following. For any element y find a power $x \in \mathbb{Z}_q$ such that $g^x = y$. The advantage of an discrete logarithm finder $\mathcal{A}$ is defined as $\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B}) = \Pr\left[\mathcal{Q}^{\mathcal{B}} = 1\right]$ where the corresponding security game is*

$$\begin{array}{l} \mathcal{Q}^{\mathcal{B}} \\ \left[\begin{array}{l} x \xleftarrow{u} \mathbb{Z}_q \\ \textbf{\textit{return}}\ [x \stackrel{?}{=} \mathcal{B}(g^x)] \end{array}\right. . \end{array}$$

1. *Show that Discrete Logarithm problem is random self-reducible, i.e., for any algorithm $\mathcal{B}$ there exists an oracle algorithm $\mathcal{A}^{\mathcal{B}}$ that for any input $y \in \mathbb{G}$ outputs the correct answer with the probability $\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})$ and has roughly the same running time.*

2. *Given that the DL problem is random self-reducible, show that the difficulty of DL instances cannot wary a lot. Namely, let $\mathcal{B}$ be a t-time algorithm that achieves maximal advantage $\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})$. What can we say about worst-case advantage*

$$\min_{y \in \mathbb{G}} \Pr\left[x \leftarrow \mathcal{A}(y) : y = g^x\right]?$$

   Can there be a large number of inputs y for which the DL problem is easy?

3. *Show how to amplify the success rate of $\mathcal{B}$ by repetitions. Sketch the corresponding time-success profile $\varepsilon(t)$. What does this say about time-success profile of DL problem in general?*

Solution. Random Self-reducibility. First, note that we can use 'pick $x \xleftarrow{u} \mathbb{Z}_q$ and set $y \leftarrow g^x$ and 'pick $y \xleftarrow{u} \mathbb{G}$' interchangeably. As g is a generator of $\mathbb{G}$, then for uniformly sampled x_0 we get that $y = g^{x_0}$ is also uniform in $\mathbb{G}$ because g generates all elements in $\mathbb{G}$ exactly once. On the other hand, for uniformly sampled y_0 there is exactly one $x_0 \in \mathbb{Z}_q$ such that $y = g^{x_0}$. Now by the assumptions of the exercise the advantage for the adversary $\mathcal{B}$ is

$$\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B}) = \Pr\left[\mathcal{Q}^{\mathcal{B}} = 1\right] = \Pr\left[x \leftarrow \mathbb{Z}_q : x = B(g^x)\right]. \tag{1}$$

However, the latter holds only if the challenge power x is uniformly sampled from $\mathbb{Z}_q$. Hence, we need to construct an algorithm $\mathcal{A}$ that feeds $\mathcal{B}$ with uniformly chosen group elements but on the same time can utilise the results. We construct the adversary $\mathcal{A}^{\mathcal{B}}$ as follows:

$$\begin{array}{l} \mathcal{A}^{\mathcal{B}}(y) \\ \left[\begin{array}{l} x_* \xleftarrow{u} \mathbb{Z}_q \\ x_{\text{guess}} \leftarrow B\left(yg^{x_*}\right) \\ \textbf{return}\ x_{\text{guess}} - x_* \end{array}\right. . \end{array}$$

First, observe that the equality $yg^{x_*} = g^x g^{x_*} = g^{x+x_*}$ implies that the event $\mathcal{B}(yg^{x_*}) = \log(yg^{x_*})$ occurs if and only if $x_{\text{guess}} - x = (x + x_*) - x_* = x_0$. In other words, the adversary $\mathcal{A}$ succeeds off $\mathcal{B}$ correctly finds the discrete logarithm of yg^{x_*} and thus

$$\Pr\left[\mathcal{A}^{\mathcal{B}}(y) = x\right] = \Pr\left[\mathcal{B}(yg^{x_*}) = \log(yg^{x_*})\right] .$$

Second, note that for any x, the element yg^{x_*} is chosen uniformly form the group. For that it sufficient to show that $x + x_*$ is uniformly sampled from $\mathbb{Z}_q$ given that x_* is uniformly sampled from $\mathbb{Z}_q$. The latter follows directly from the equivalence $x_0 = x + x_* \Leftrightarrow x_* = x_0 - x$, which formally assures

$$\Pr\left[x_* \leftarrow \mathbb{Z}_q : x + x_* = x_0\right] = \Pr\left[x_* \leftarrow \mathbb{Z}_q : x_* = x_0 - x\right] = \frac{1}{|\mathbb{G}|} .$$

Combining these two facts, we get

$$\Pr\left[\mathcal{A}^{\mathcal{B}}(y) = x\right] = \Pr\left[\mathcal{B}(yg^{x_*}) = \log(yg^{x_*})\right] = \Pr\left[y \leftarrow \mathbb{G} : \mathcal{B}(y) = \log y\right] = \mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B}) \ .$$

Finally, let us estimate the running-time of $\mathcal{A}^{\mathcal{B}}$ by listing all operations

- uniform sampling from $\mathbb{Z}_q$;
- exponentiation g^{x_*};
- multiplication yg^{x_*};
- call to adversary $\mathcal{B}$;
- subtraction $x_{\text{guess}} - x_*$.

The time for all operations except the call to the adversary $\mathcal{B}$ depends on the group size $|\mathbb{G}|$. If the group $\mathbb{G}$ is fixed, then we can denote this overhead with δ. So the total time required for the new adversary is $t + \delta$. Hence, have constructed a $(t + \delta)$-time adversary $\mathcal{A}^{\mathcal{B}}$ which has advantage $\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})$ for any input $y \in \mathbb{G}$.

Uniformity. So far we have established that for any fixed $y \in \mathbb{G}$ probability $\Pr\left[x \leftarrow \mathcal{A}(y) : y = g^x\right] = \mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})$ holds. Consequently,

$$\min_{y \in \mathbb{G}} \Pr\left[x \leftarrow \mathcal{A}(y) : y = g^x\right] = \min_{y \in \mathbb{G}} \mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B}) = \mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B}).$$

Informally, there are no *hard instances* which are more difficult to compute, as we can always mask then as randomly chosen instances with a small computational overhead δ. Similarly, there cannot be many easy instances of discrete logarithm unless the discrete logarithm problem is really easy. For example, let

$$\mathcal{E}(\kappa) = \{y \in \mathbb{G} : \Pr\left[x \leftarrow \mathcal{B}(y) : g^x = y\right] > \kappa\}$$

be the set of easy instances and let

$$\mathcal{H}(\kappa) = \{y \in \mathbb{G} : \Pr\left[x \leftarrow \mathcal{B}(y) : g^x = y\right] \leq \kappa\}$$

be the complementary set of difficult instances. Then we can express the average success against discrete logarithm as follows

$$\begin{aligned}
\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B}) = \Pr\left[x \xleftarrow{u} \mathbb{Z}_q : \mathcal{B}(g^x) = y\right] &= \frac{1}{q} \cdot \sum_{y \in \mathbb{G}} \Pr\left[\mathcal{B}(y) = \log(y)\right] \\
&= \frac{1}{q} \cdot \sum_{y \in \mathcal{E}(\kappa)} \Pr\left[\mathcal{B}(y) = \log(y)\right] + \frac{1}{q} \cdot \sum_{y \in \mathcal{H}(\kappa)} \Pr\left[\mathcal{B}(y) = \log(y)\right] \\
&> \frac{1}{q} \cdot \sum_{y \in \mathcal{E}(\kappa)} \kappa = \frac{\kappa}{q} \cdot |\mathcal{E}(\kappa)| \ .
\end{aligned}$$

Hence, the fraction of easy instances is given by the ratio between κ and $\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})$:

$$|\mathcal{E}(\kappa)| < \frac{\mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})}{\kappa} \cdot q \ .$$

For instance, the set instances where the discrete logarithm is more than 1000 times simpler than on average can be at most a thousandth of the entire group. Hence, if the large set of easy instances, then the discrete logarithm problem itself is also easy. To put it another way, if the set of easy instances is large then the

random self-reduction recasts the challenge into the easy instance with high probability—discrete logarithm must be easily solvable.

Amplification effects. To amplify the advantage of $\mathcal{B}$, we can contract the following adversary

$$
\begin{array}{l}
\mathcal{C}^{\mathcal{B}}(y, t_{\max}) \\
\left[\begin{array}{l}
\text{while runningtime is smaller than } t_{\max}: \\
\quad \left[\begin{array}{l}
x_* \xleftarrow{u} \mathbb{Z}_q \\
x_{\text{guess}} \leftarrow \mathcal{B}(y g^{x_*}) \\
\text{if } g^{x_{\text{guess}} - x_*} = y \text{ then } \quad \textbf{return } x_{\text{guess}} - x_* \ .
\end{array}\right.
\end{array}\right.
\end{array}
$$

As each loop takes roughly time t, the number of iterations is approximately $k \approx t_{\max}/t$. For brevity, let us denote $\varepsilon = \mathsf{Adv}^{\mathsf{dl}}_{\mathbb{G}}(\mathcal{B})$. Then, in every loop, the probability that the adversary continues is $1 - \varepsilon$ and that it stops is ε. Thus, the failure rate after $k \approx t_{\max}/t$ loops is $(1-\varepsilon)^k$. So the success rate is $1 - (1-\varepsilon)^k$, which is approximately $k\varepsilon$ whenever the overall success probability is below 10%. Indeed, if $\varepsilon k \leq 10\%$ then the contribution of secondary terms to the binomial sum hardly noticeable, as

$$\frac{\binom{k}{2}\varepsilon^2}{\binom{k}{1}\varepsilon} = \frac{(k-1)\varepsilon}{2} \leq 5\% \ .$$

The contribution of higher order terms is orders of magnitudes less. Thus, the increase in success is quasi-linear with relative error around $5 - 10\%$, as illustrated in the following figure.