Solution

1. **(4 pts.) Any questions**
   Any constructive responses is given full credit.

2. **(20 pts.) PGP**
   If you emailed your TA with a correctly signed encrypted message you will receive full credit.

3. **(10 pts.) One-time pad**

   (a) No, this scheme does not have the security guarantees of a one-time pad. Table 1 lists the resulting encrypted messages using this scheme. We can see that some outcomes exclude certain inputs. For example, given $(M,K) = 11$ an attacker knows that the sent message M is not 0.

   (b) We wish to design a new encryption algorithm $E^*(\cdot,\cdot)$ that has the security guarantees of the one-time pad. We require that given $E^*(M,K)$, an attacker should get no information about M. This property is satisfied for any $E^*(M,K)$ that is uniform on $\{0,1,2\}$. One such algorithm is as follows:

   $$E^*(M,K) = M + K \text{ mod } 3.$$

   Table 2 confirms that each outcome is equally likely.

4. **(10 pts.) An RSA reduction**
   We wish to factor $N = pq$. Since $e = 3$ and d are inverses modulo $\varphi(N) = (p-1)(q-1)$, have that

   $$3d = ed = 1 + k(p-1)(q-1) = 1 + k\varphi(N)$$

   for some $k \in \{1,2,\dots\}$. Also we have that $d < \varphi(N)$, so $k \in \{1,2\}$. (In fact $k = 2$ always.)

Table 1: Encrypted messages using E

M	K	E(M,K)
00	00	00
00	01	01
00	10	10
01	00	01
01	01	00
01	10	11
10	00	10
10	01	11
10	10	00

Table 2: Encrypted messages using E^*

M	K	$E^*(M,K)$
00	00	00
00	01	01
00	10	10
01	00	01
01	01	10
01	10	00
10	00	10
10	01	00
10	10	01

We have a finite number of possible values of k, so we can check which k is correct as follows:
Fix a k. Given this guess at k, we can infer a presumed values for $\varphi(N)$ via

$$\varphi(N)_k = \frac{3d-1}{k}.$$

Also the true value for $\varphi(N)$ satisfies $\varphi(N) = (p-1)(q-1) = pq - p - q + 1 = N - \frac{N}{q} - q + 1$; rewriting this, we can solve for q via the quadratic equation, given the value of $\varphi(N)$. This gives a way to test whether our guess $\varphi(N)_k$ was correct, since we can use our guess to solve for q and test whether the resulting q is indeed a factor of N.

The running time is polynomial in the number of bits of N: we use $O(1)$ operations on integers no larger than N, which corresponds to $O((\lg N)^2)$ bit operations.

An algorithm for general e is given in G. Miller, "Riemann's hypothesis and tests for primality," *Journal of Computer and System Sciences*, 13(3):300-317, 1976. This algorithm is in time polynomial to the number of bits in N.

5. (21 pts.) The definition of a secure block cipher

(a) Insecure
A distinguishing attack on the block cipher E is as follows:
Ask for the encryption of two messages M, M'; receive the ciphertexts C, C'. If $M \oplus C = M' \oplus C'$, then guess that you are interacting with E; otherwise, you are definitely interacting with P. This works because $M \oplus E_K(M) = K$ for all M, yet the corresponding equality occurs with extremely low probability for P. The distinguishing advantage of this attack is $1 - 1/(2^{128} - 1)$.

(b) Secure
A reduction proving the security of E goes as follows:
Suppose there is some successful distinguishing attack A that breaks E. Define the attack B by $B^f = A^{f(\cdot)\oplus\mathbf{1}}$. In other words, B simulates the operation of A, except that when A hands message M to its box, B queries its box with M, receives C, hands $C \oplus \mathbf{1}$ back to A (as though it were the response from A's box), and continues to simulate A. I claim that B breaks AES. In particular, $B^{\mathrm{AES}_K(\cdot)} = A^{E_K}$, and $B^{P(\cdot)} = A^{P'(\cdot)}$, where $P'(x) = P(x) \oplus \mathbf{1}$. Now if P is a random permutation, then so is P'. This means that B distinguishes AES from a random permutation, with advantage $\mathrm{Adv}\, B = \mathrm{Adv}\, A$. In summary, if there is any attack that breaks E (distinguishes E from random with advantage ε), then there is an attack that breaks AES (distinguishes AES from random with

the same advantage ε). Taking the contrapositive, we see that if AES is secure (there is no attack that breaks AES), then E is secure (there is no attack that breaks E).

(c) Insecure
An attack on the block cipher E is as follows:
Ask for the encryption of message M, receiving C; if $C = \text{AES}_{\mathbf{0}}(M)$, then guess that you are interacting with E; otherwise, you are definitely interacting with P. This attack has advantage $1 - 1/2^{128}$ at distinguishing E from P.

(d) Insecure
An attack on the block cipher E is as follows:
The attack is a combination from part (a) and (c). Ask for the encryption of two messages M, M'; receive the ciphertexts C, C'. Decrypt C and C' under AES with key $\mathbf{0}$. If $M \oplus \text{AES}_{\mathbf{0}}^{-1}(C) = M' \oplus \text{AES}_{\mathbf{0}}^{-1}(C')$, then guess that you are interacting with E; otherwise, you are definitely interacting with P.

(e) Secure
Here is a sketch of a proof:
Since AES is secure, you cannot distinguish $\text{AES}_{K_1}(\text{AES}_{K_0}(\cdot))$ from $\text{AES}_{K_1}(P(\cdot))$. (One can prove this by a reduction: if A distinguishes these two, then B, given by $B^f = A^{\text{AES}_k(f(\cdot))}$ where k is chosen randomly, distinguishes AES from P, which is impossible.) Also, since AES is secure, you cannot distinguish $\text{AES}_{K_1}(P(\cdot))$ from $P'(P(\cdot))$, where P' is a random permutation chosen uniformly and independently at random from P. (Another reduction: $B^f = A^{f(P(\cdot))}$.) But $P' \circ P$ is also a random permutation. Combining the above statements, we see that $E_K(\cdot)$ cannot be distinguished from a random permutation.

(f) Insecure
An attack:
Let $M = (M_L, M_R)$ be arbitrary. Choose $M' = (M_L, M'_R)$, so that M and M' have the same left half. Ask for the encryption of M and M', receiving C and C'. Check whether C and C' agree in their left half. If $C_L = C'_L$, guess that you are interacting with E; otherwise, you are definitely interacting with P.

(g) Insecure
An attack:
Choose M arbitrarily. Ask for the encryption of M, receiving C. Ask for the encryption of C, receiving C'. Now if $C' = M$, guess that you are interacting with E; otherwise, you are interacting with P. This works since

$$\begin{aligned} E_K(E_K(M)) &= \text{AES}_K^{-1}(\text{AES}_K(\text{AES}_K^{-1}(\text{AES}_K(M) \oplus \mathbf{1})) \oplus \mathbf{1}) \\ &= \text{AES}_K^{-1}((\text{AES}_K(M) \oplus \mathbf{1}) \oplus \mathbf{1}) \\ &= \text{AES}_K^{-1}(\text{AES}_K(M)) \\ &= M \end{aligned}$$

but $P(P(M))$ is rarely equal to M.

6. (35 pts.) Security of CBC encryption

(a) If the inputs to the cipher P are distinct, then the distribution on the ciphertext $C = (C_0, \ldots, C_j)$ is as follows: it is uniformly distributed on the set of ciphertexts such that all the C_i's are distinct. Thus the distribution of the cipher text C when box is type I is the same as when the box is type II, if P is invoked on distinct inputs.

Since the distribution of the cipher text C is the same for both type I and type II box, the adversary has no advantage in distinguishing between the encryption of M and the encryption of M'. Formally, the distribution on the output of $A^{\text{box of type I}}$ is the same as the distribution on the output of $A^{\text{box of type II}}$ (since everything A sees has the same distribution regardless of the type of the box), so $\Pr[A^{\text{box of type I}} = \text{type I}] = \Pr[A^{\text{box of type II}} = \text{type I}]$, and thus $\mathrm{Adv} A = 0$.

(b) Let E_m denote the event that $D_0, \ldots, D_{m-1}$ are all distinct. We wish to prove that $\Pr[D_m = D_i | \mathsf{E}_m] \leq 1/(2^n - m)$.

Note that $D_m = C_{m-1} \oplus B_m = P(D_{m-1}) \oplus B_m$. Therefore, $D_m = D_i$ holds if and only if $P(D_{m-1}) \oplus B_m = D_i$, or equivalently, if and only if $P(D_{m-1}) = D_i \oplus B_m$. This means our goal is to show that $\Pr[P(D_{m-1}) = D_i \oplus B_m | \mathsf{E}_m] \leq 1/(2^n - m)$.

Suppose $D_0, \ldots, D_{m-1}$ and $P(D_0), \ldots, P(D_{m-2})$ have all been chosen. Then $D_i \oplus B_m$ is just a fixed bit string. Also (by assumption) D_{m-1} is different from all of $D_0, \ldots, D_{m-2}$. Thus the distribution on $P(D_{m-1})$ is uniform on the set of all values other than $P(D_0), \ldots, P(D_{m-2})$. In particular, the value that $P(D_{m-1})$ hits any fixed bit string in this case is $\leq 1/(2^n - m)$.

(c) We want to bound the probability that $D_i = D_j$ for some $i \neq j$. Let E denote the event that all D_i's are distinct, so we are looking for an upper bound on $\Pr[\neg \mathsf{E}]$. Define A_m as the event that $D_m = D_i$ for some $i < m$. By part (b),

$$\Pr[A_m | \neg(A_1 \cup \cdots \cup A_{m-1})] \leq \sum_{i=0}^{m-1} \Pr[D_m = D_i | \mathsf{E}_m] \leq \sum_{i=0}^{m-1} 1/(2^n - m) = m/(2^n - m).$$

Now we are ready to calculate the desired bound:

$$\begin{aligned}
\Pr[\neg \mathsf{E}] &= \Pr[A_1 \cup A_2 \cup \ldots \cup A_j] \\
&\leq \Pr[A_1] + \Pr[A_2 | \neg A_1] + \cdots + \Pr[A_j | \neg(A_1 \cup \cdots \cup A_{j-1})] \\
&\leq \sum_{k=1}^{j} \frac{k}{2^n - k} \qquad \text{(see above)} \\
&\leq \frac{1}{2^n - j} \sum_{k=1}^{j} k \\
&= \frac{1}{2^n - j} \times \frac{j(j+1)}{2} \\
&= \binom{j+1}{2} \frac{1}{2^n - j}.
\end{aligned}$$

(d) First we answer the questions from the hint. Event E is as above—namely, when all the inputs to cipher P are unique. By part (c), $\Pr[\mathsf{E}] \geq 1 - \binom{j+1}{2}/(2^n - j)$. Also $\Pr[A^{\text{box of type I}} = \text{type I} | \mathsf{E}] = \Pr[A^{\text{box of type II}} = \text{type I} | \mathsf{E}]$ (by part (a)).

Now we can calculate the advantage of the adversary at breaking CBC-P:

$$\begin{aligned}
\text{Adv}A &= \left|\Pr[A^{\text{box of type I}} = \text{type I}] - \Pr[A^{\text{box of type II}} = \text{type I}]\right| \\
&= \big|\Pr[\neg\mathsf{E}]\cdot\Pr[A^{\text{box of type I}} = \text{type I}|\neg\mathsf{E}] + \Pr[\mathsf{E}]\cdot\Pr[A^{\text{box of type I}} = \text{type I}|\mathsf{E}] \\
&\quad - \Pr[\neg\mathsf{E}]\cdot\Pr[A^{\text{box of type II}} = \text{type I}|\neg\mathsf{E}] - \Pr[\mathsf{E}]\cdot\Pr[A^{\text{box of type II}} = \text{type I}|\mathsf{E}])\big| \\
&= \big|\Pr[\neg\mathsf{E}]\cdot\Pr[A^{\text{box of type I}} = \text{type I}|\neg\mathsf{E}] - \Pr[\neg\mathsf{E}]\cdot\Pr[A^{\text{box of type II}} = \text{type I}|\neg\mathsf{E}] \\
&\quad + \Pr[\mathsf{E}]\cdot\Pr[A^{\text{box of type I}} = \text{type I}|\mathsf{E}] - \Pr[\mathsf{E}]\cdot\Pr[A^{\text{box of type II}} = \text{type I}|\mathsf{E}])\big| \\
&= \left|\Pr[\neg\mathsf{E}]\cdot\Pr[A^{\text{box of type I}} = \text{type I}|\neg\mathsf{E}] - \Pr[\neg\mathsf{E}]\cdot\Pr[A^{\text{box of type II}} = \text{type I}|\neg\mathsf{E}]\right| \\
&\qquad\qquad \text{(the last two terms were equal, by part (a))} \\
&= \Pr[\neg\mathsf{E}]\cdot\left|\Pr[A^{\text{box of type I}} = \text{type I}|\neg\mathsf{E}] - \Pr[A^{\text{box of type II}} = \text{type I}|\neg\mathsf{E}]\right| \\
&\le \Pr[\neg\mathsf{E}] \qquad\qquad \text{(since } |p-q| \le 1 \text{ whenever } 0 \le p,q \le 1\text{)} \\
&\le 2\binom{j+1}{2}/(2^n - j). \qquad\qquad \text{(by part (d))}
\end{aligned}$$

(e) Recall that box I computes the function $(M,M') \mapsto$ CBC-AES$_K(M)$, and box II computes $(M,M') \mapsto$ CBC-AES$_K(M')$. Our goal is to show that these two boxes are indistinguishable.

- Let box I$'$ represent the function $(M,M') \mapsto$ CBC-$P(M)$. We will first show that box I is indistinguishable from box I$'$.
  Proof: By a reduction. Suppose A is an attacker that distinguishes box I from box I$'$. We'll define an algorithm B that distinguishes a AES$_K(\cdot)$ box from a $P(\cdot)$ box. B^f works by simulating A and using its (B's) box f to emulate CBC-$f(\cdot)$; this is possible, since CBC only uses AES or P as subroutines. If A sends (M,M') to its box, where $M = (M_1,\ldots,M_j)$, then B will pick a random IV D_0, compute $C_0 = f(D_0)$ and $C_i = f(C_{i-1} \oplus M_i)$ for $i = 1,\ldots,j$, and return $C = (C_0,\ldots,C_j)$ to A. Finally, B outputs whatever A does. Note that $B^{AES_K(\cdot)} = A^{\text{box I}}$ and $B^{P(\cdot)} = A^{\text{box I}'}$. Therefore Adv$B$ = Adv$A$; but by assumption, we know that Adv$B \le T/2^l$. In conclusion, there is no way to distinguish box I from box I$'$ with advantage greater $T/2^l$.
- Let box II$'$ represent the function $(M,M') \mapsto$ CBC-$P(M')$. By a very similar argument, box II cannot be distinguished from box II$'$, except with advantage $\le T/2^l$.
- Finally, by part (e), box I$'$ cannot be distinguished from box II$'$, except with advantage $\le 2\binom{j+1}{2}/(2^n - j)$.

This means that box I cannot be distinguished from box II except with advantage $\le 2T/2^l + 2\binom{j+1}{2}/(2^n - j)$.