ECE 750: Usable Security and Privacy - Final	Marks obtained ↓
Date: Aug 7, 2025 Total questions: **4** Total points: **50**	
ID: Name: Time: 2.5 hrs	

Page:	3	4	6	7	8	9	10	11	Total
Points:	9	7	7	6	4	3	10	4	50
Score:									

Instructions

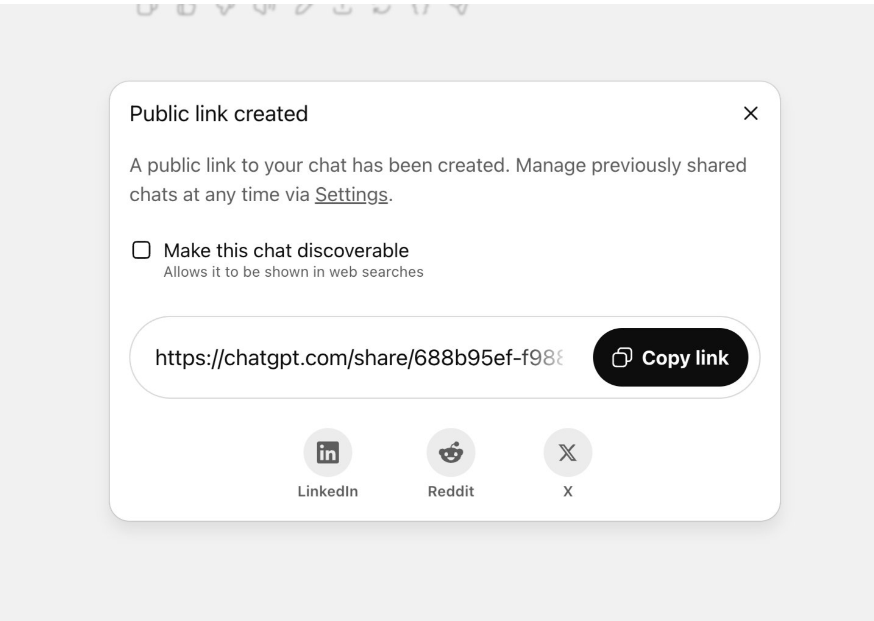
No aids allowed. All you are allowed is a pen and pencil.

Use space provided. Answer the questions in the spaces provided. If you run out of room for an answer extra pages are provided at the end of the test booklet starting on page 12. They are clearly marked as EXTRA ANSWER SPACE. Please state in the original answer space if the extra pages are being used so that the grader knows to look there.

Point value in right-hand margin. The number of points each question is worth is listed in the right hand margin. The number of points and the space provided are roughly indicative of how long of an answer is expected.

Pencils and pens allowed. Both pens and pencils are allowed. Pencil marks need to be clearly present or erased. If it is unclear if a mark is or is not present then it is up to the discretion of the grader and this point cannot be contested later.

Solution: This version of the exam has sample solutions. Most of the exam questions accept multiple possible answers. The answers in the Solution box are the answer the instructor had in mind when the question was written.



Public link created

A public link to your chat has been created. Manage previously shared chats at any time via Settings.

□ Make this chat discoverable
Allows it to be shown in web searches

https://chatgpt.com/share/688b95ef-f98

Figure 1: ChatGPT screenshot of a link to the chat conversation being created. User has the option to make the chat "discoverable". Screenshot is from arsTechnica

ChatGPT "Discoverable" Chats

1. Users of ChatGPT may want to share their chat conversations with others so ChatGPT offers a feature where users can click "share" on the main chat interface and get a URL link they can share with others. OpenAI, which operates ChatGPT, recently decided to test out a new feature where users can not only share their chats via link, they can also make them discoverable to search engines, thereby making the chats public to anyone. To preserve privacy, the feature was opt-in. To share the chat's with search engines a user had to tick the "Make this chat discoverable" box shown in Figure 1.

   We now know that many users likely ticked the opt-in box in error causing their obviously-personal chats to become indexed by search engines like Google and Bing. The chats were shared anonymously, but the content of the chat itself could still expose identity. For example, someone asked ChatGPT to help with their resume and provided the resume file in the chat. The resume contained their name and therefore connected the chat to them even though the username was anonymised.

   Assuming making chats discoverable to search engines was in error, here must be something about the interface caused users to tick the "make this chat discoverable" box when they did not want their chat to be discoverable.

   Answer the following questions about the ChatGPT discoverable chats.

(a) A common argument against strong privacy is that "If you have nothing to hide, you have nothing to worry about." Thinking about what we have learned about the concept of privacy in this course, provide a response to this argument in the context of ChatGPT chats being made discoverable via search engines. (5)

Solution: Many possible answers.

The "If you have nothing to hide, you have nothing to worry about." argument presumes that the only harm a person might experience has to do with them doing something wrong. But this is a flawed argument as there are many ways a person could experience negative consequences without doing anything wrong. A simple example is negotiating the price of a used car where there is already an imbalance in knowledge between the salesperson and the buyer. Providing the salesperson with even more data such as income of the buyer will only hurt the buyer, even though they have done nothing wrong.

Chilling effects are another argument. If a person feels that their privacy might be violated through engagement, they may choose to not engage at all. This is called a chilling effect and it is generally bad for society as it means that people may choose to not speak out, not sign petitions, or even not purchase items which may impact GDP.

(b) Would a cognitive walkthrough be an appropriate way to find the usability issue? State 'yes' or 'no' and then explain your reasoning. (4)

Solution: No. A cognitive Walkthrough is done based on "correct" steps and focuses on if users can do those steps. This problem is caused by users taking an incorrect step. The action of clicking the tick box is unlikely to be listed as an action on the walkthrough.

Yes is possible, though a less good answer. An informal cognitive walkthrough might notice the issue. But still, while the methodology can find potential errors a user might make, the main focus of the methodology is understanding barriers users have to completing the correct set of actions.

(c) Bera thinks the issue may be the words "discoverable" and "web searches". She thinks that end-users and web developers have different interpretations of these words in the context in which they are shown. (7)

To test her hypothesis she designed an online survey with the following structure:

- Participants were shown informed consent and consented
- They are then randomly assigned to one of two groups.

  Group A: is shown the interface in Figure 1 with the "Make this chat discoverable" tick box as shown.

  Group B: is shown an interface similar to Figure 1 but the tick box instead says: "Turn on public access" in larger text and "Allows anyone to see your chat" in smaller light gray font designed to follow the same font sizes and styles as Figure 1.
- Participants are asked: "Would you tick the tick-box pictured above when sharing a ChatGPT chat." with 'yes' and 'no' being the options.
- Participants are given an open text box and asked "What do you think the text next to the tick-box means?"
- Participants are finally asked their: gender, nationality, age in years, and what languages they are proficient in.
- Participants are thanked and the survey ends.

Identify two design flaws in this study that cause serious internal validity problems.

Solution: The largest design flaw is that it does not separate developers from other users. This study will at best tell us what normal users think the language means. But it will not tell us if developers and end-users interpret the words differently.

The study design also does not target the wording. It does show different wording, but the wording is so different that after the study is run we will still not know if the word "discoverable" was the problem or if it was one of the other words in the message. Or if the new message is just much clearer.

Context is also almost completely missing. Users are only asked if they would tick when "sharing" but who is being shared with is not explained.

QUESTION 1.c ANSWER SPACE

Passwords

2. (a) Both entropy and guessability are metrics used to measure password strength. How are the concepts of *entropy* and *guessability* different? (5)

Solution: Entropy measures the potential size of the password space for a given password. In other words: if an attacker were to brute-force try every password combination from the set of unique possible symbols, how many guesses would be required to try all combinations?

Guessability measures how long it would take an attacker to guess a password assuming that the attacker is starting with a dictionary of known prior passwords people have used in the past. Assuming that commonly-used passwords will be tried first, how many guesses would it take to find a given password?

The two approaches are different in several ways:

- Entropy can be calculated just from knowing the length and full character set. So it is easy to compute.
- Guessability requires knowledge and assumptions about how the attacker will approach guessing, so it has more variance in how it is computed.
- Guessability will give a poor score to a password like: `P@$$w0rd123` but entropy will give it a high score because it is long (11 characters) and uses a large character set.

(b) State an example password that would have a high entropy but be considered very guessable. (2)

Solution: Any password that is drawn from a large character set but is actually very common. `P@$$w0rd123`

Methodology

3. The below quote is from the paper *"... no one Can Hack My Mind": Comparing Expert and Non-Expert Security Practices by Ion, Iulia, Rob Reeder, and Sunny Consolvo*:

> To design the surveys, we first conducted in-person semi-structured interviews of 40 security experts at the 2013 BlackHat, DefCon, and USENIX Security conferences. We defined experts as conference attendees who reported having at least 5 years of experience working in or studying computer security. We started every interview with our top-3-advice question:
>
> *What are the top 3 pieces of advice you would give to a non-tech-savvy user to protect their security online?*
>
> We asked follow-up questions to clarify responses. Interviews lasted 8 minutes on average. We transcribed all of the interviews and coded the advice we collected. Interview data informed many of the questions we asked in the surveys, as we note below.

The paper used a 2-part methodology. The first part (described in the quote above) is an interview of 40 people. The results of the interview were then used to create two similar surveys: A survey of experts (n=231) and a survey of non-experts (n=294).

(a) The aim of this sequence of studies is to understand security practices of normal users. Describe how conducting an interview study first might improve the *external validity* of the later survey design. (6)

Solution: The later survey is likely going to need a finite list of advice for users to choose between. Or at least the survey designers are going to need to know what possible types of advice people are likely to give and if that advice is "good". By studying expert-advice the survey can be designed such that the advice options match what experts think.

(b) The researchers in the above study found that experts recommend using password managers but they suspect that end users do not consider password managers very effective. Imagine that they decided to test this theory using a survey. The survey contained the following question: (4)

State how much you agree or disagree with the following statement:
Password managers store passwords safely.

Participants were given a 5-point Likert scale with the options ranging from “Strongly agree” to “Strongly disagree”.

During analysis, the researchers divided the participants into groups based on a demographics question about if they had any prior software development experience or not. They then ran a two-sample t-test on the answers to the password manager question to see if the two groups have different means.

p-value: .00003

α**:** .05

Software developer mean: 4.4

Not a software developer mean: 3.9

Interpret the result. Is there a difference in the two groups? What does the result mean in terms of the question the survey aimed to answer.

Solution: There is a statistically significant difference between the two groups. Since people with software development experience have a mean of 4.4 (Strongly Agree = 5), they are more likely to agree that password managers store passwords safely than non-software-develper users.

(c) Based on interview described at the top of Questions 3 and the t-test from Question 3.b, mark each of the following statements as proven to be true (T) or not yet proven (F). (3)

Please mark **T** for proven true, and **F** for not currently proven.

______ Experts as defined in the interview study at the top of Question 3 are more confident in password manager safety than general internet users.

Solution: F and T were both accepted. Originally intended to be F, but the interview study could also be seen as "proof", so both answers accepted.

______ People with software development experience consider password managers to be safer than people with no software development experience.

Solution: T - This is literally what the t-test is testing. So it should be True.

______ Password managers are safer than writing passwords down.

Solution: F - Neither study tests this statement.

Public/Private Key Cryptography

4. (a) Fill in the blanks in the following text describing Alice and Bob correctly communicating securely. Each blank needs to be filled in using one of the following words. The words can be used more than once and not all words need to be used: (6)

- Encrypt, encrypted, encrypts
- Decrypt, decrypted, decrypts
- Sign, signed, signs, signature
- Public
- Private

When they last met in person Alice and Bob verified and then __________ each other's __________ keys using their respective __________ keys.

Later after they have gone home, Alice decides to send an encrypted email to Bob. Alice first __________ the message using her __________ key and then __________ the resulting message using Bob's __________ key. Alice then sends the message normally using a potentially untrusted connection.

Bob receives the message and first __________ it using his __________ key. He then verifies that the message really was from Alice by verifying the __________ using Alice's __________ key.

Solution: Alice wants to send an encrypted and signed message to Bob who she has met in the past. When they last met in person they verified and then **signed** each other's **public** keys using their respective **private** keys.

To send an email to Bob, Alice first **signs** the message using her **private** key and then **encrypts** the resulting message using Bob's **public** key. Alice then sends the message over an untrusted connection.

Bob receives the message and first **decrypts** it using his **private** key. He then verifies that the message really was from Alice by verifying the **signature** using Alice's **public** key.

(b) Certificate Authorities are intended to make everyday encryption between end users and websites more usable. How does using Certificate Authorities make encryption more usable than the interaction described in Question 4.a? Assume the user has reasonably-designed software helping them, but that the software is not hiding important steps in the protocol. (4)

Solution: A CA replaces the step where Alice and Bob have to sign each other's public keys using their private keys. Instead a CA certifies their identity on their behalf. Using a CA makes the process more usable because users do not have to individually verify each person they interact with, they only need to trust the relevent CA to do the verification.

(c) One of the reasons research in Usable Security and Privacy is so challenging is the “barn door” property which refers to the fact that some user actions are not possible to reverse and that “closing the barn door *after* the horse has escaped” is not very effective. (4)

Describe two errors that a user could make while trying to correctly engage in the above Alice and Bob interaction that are irreversible and cannot be recovered from. Think about some of errors you read about in Why Johnny Can’t Encrypt.

Solution: Some examples:

- Confusing public and private keys. Encrypting using a private key instead of the other person’s public one makes the message world-readable to anyone with the public key which is theoretically everyone.
- Signing without encrypting. This will guarantee integrity but not confidentiality. The message is now public and cannot be taken back.
- Emailing the private key to the other person or making the private key public in any way.
- Losing the private key. While “safe” in that no data will be disclosed, this error is not possible to recover from and all messages encrypted with the paired public key become non-recoverable.

Extra Answer Space

If you need extra space to give an answer, please state in the original answer space that you will be using the extra pages and continue or write your answer here. If you choose to use the extra space as scratch paper, please write “scratch” so that graders know to ignore anything you have written.

EXTRA ANSWER SPACE