FIT2093 Introduction to Cybersecurity - 2023

Assignment 2: User Authentication and Access Control

Purpose	*User authentication and access control are important aspects in cybersecurity. In the first part of this assignment, you will apply the concepts we learned on user authentication into practice to assess two widely-used user authentication approaches, and their security & usability aspects. Once users are authenticated to a system, the system gives them authorization to access system resources based on their access rights. In the second part of this assignment, you will put your knowledge on access control into practice to analyse the security of a multi user file system and apply appropriate access control configurations to achieve its security requirements*
Your task	*This assignment is an individual assessment. You are given tasks on two widely-used authentication approaches: biometric and password, as well as the practical configuration of access control in a multiuser system.*
Value	**20**% of your total marks for the unit
Page / Time Limit	**ONE** Individual video**:** Presentation **up to 15 minutes** **ONE** Individual pdf: Report up to **10 pages** (excluding cover page and appendix)
Due Date	**9 May 2023 11:55 pm Melbourne time**
Submission	Submit two files for all tasks: ● Individual Report for **all tasks** as a **pdf** via Moodle assignment submission. ● Individual Video for **all tasks** as **mp4** via Moodle assignment submission. ● Turnitin will be used for similarity checking of all written submissions.
Late Penalties	● 10% deduction per calendar day or part thereof for up to one week ● Submissions more than 7 calendar days after the due date will receive a mark of zero (0) and no assessment feedback will be provided.
Feedback	Feedback will be provided on student work via: ● general cohort performance ● specific student feedback ten working days post submission

Background

Multiple factor authentication, usually a combination of biometric and password authentication factors, is widely used today to verify the identity of a user attempting to access a system. After being verified, whether they are permitted to access certain resources in the system depends on the access control configured. This assignment is designed to improve your understanding of these topics and apply this understanding in a practical real world scenario.

Task 1 gives you the practical exercises on a two-factor authentication system while Task 2 is an access control exercise of how a network administrator can set the permissions to various users and user groups according to the system access control requirements.

Overview of the assignment

The assignment is worth 20% of your total unit mark. In **Task 1** of the assignment (weight: 10% of your unit mark), you are to prepare an individual video to demonstrate your understanding of biometric authentication and password authentication.
For biometric authentication, you will use a given Face Recognition system, CompreFace (**Option 1**), or a sample face recognition testing data set (**Option 2**), to study how the system parameters affect the authentication accuracy. For password authentication, you will use the UNIX password authentication and use an attack tool application, John the Ripper and the UNIX password hashing tools, to study the impacts of password hashing parameters on the difficulty of hacking the hashed password and the usability of the system.

In **Task 2** of the assignment (weight : 10% of your unit mark), you are given several practical tasks on the access control knowledge you acquired in the lectures and tutorials/labs. You are asked to create new users and assign them into appropriate user groups, as well as study how SUID is used to give special access rights to users running certain processes.

An **individual report** and a **video** containing all tasks in this assignment have to be submitted via Moodle links. In the report, you have to illustrate your results and the explanations for each task. You will also prepare a video of up to 15-minute presentation for each task.

To complete the tasks in this assignment, you have two options and you can choose to submit either option:

- **Option 1 (using CompreFace and original Asg2 VM)**: This is the original option and is identical to the one specified in the earlier version (ver 1.0) of this assignment spec doc, **except for a couple of corrections to the Task 1b (indicated in purple font)**. For this option, a special VM is required to be installed on your computer containing the CompreFace face recognition software and is available for download from the Moodle Assessments page (note that this VM is different from the FIT2093 Lab VM). As the VM takes resources on computation, you are required to adjust the settings recommended in the Appendix.

- **Option 2 (using a sample face recognition testing data set for Task 1a and the Asg2Option2 VM for Task 1b and Task 2)**: For this option, for Task 1a, you will use

a sample face recognition testing data set provided in the Assignment Details Section of this document for Task 1a Option 2. For this option, you DO NOT need to use the CompreFace face recognition software for Task 1a. For Task 1b and Task 2 with this option, you can use the "**FIT2093_Asg2_LW.ova" VM** available for download from the Moodle `Assessment' page. This VM is lightweight and similar in system requirements to the FIT2093 Lab VM, so you should be able to install it on your device similarly to the FIT2093 Lab VM.

NOTE: If you choose this option, add the label "OPTION 2" to your submitted report/video front page.

Assessment Details

Task	Rubric
Task 1a	5% (2.5% from video and 2.5% from report) ■ 2% for the numerical answers (FAR and FRR) of each threshold ■ 2% for the discussion on the significance of the threshold ■ 1% for the choice of threshold
Task 1b	5% (2.5% from video and 2.5% from report) ■ Results of the four hashed passwords: 0.25% for each (0.25%x8) including screen captures for both john and mkpasswd timing. ■ Discussions on the different approaches (1.5%) and time estimates for dictionary search (0.5%) ■ Recommendations on the password hashing (1%)
Task 2	10% (5% from video and 5% from report) ■ Subtask 2a (4%: 1% of each subparts (1) - (4)). ■ Subtask 2b (3%: 1% on configuration and 2% read the file by non-owner) ■ Subtask 2c: (3%: 1% vulnerability test and 2% explanation and mitigation if it exists)

Assignment Details

Task 1 (10% of unit marks): Two-Factor Authentication System

Task 1a) (5% of unit marks): Biometric authentication

In this task,you have two options and you can choose to submit either option. Please follow the corresponding instructions below for your chosen option.

- **Option 1 (using CompreFace and original Asg2 VM)**: This option for Task 1a is identical to the Task 1a specified in the earlier version (ver 1.0) of this assignment spec doc. An open source software application for face identification, CompreFace, is used for computing the biometric authentication factor. CompreFace and face collections of over 5000 people have been pre-installed in the Assignment 2 VM.

- **Option 2 (using a sample face recognition testing data set)**: For this option, you will use a sample face recognition testing data set provided in the instructions for Task 1a Option 2 below. For completing Task 1a with Option 2, you DO NOT need to use the CompreFace face recognition software or any other VM.

Using your own testing and given face image samples (Option 1) or the sample data sets (Option 2), you are asked to evaluate the security and usability of this biometric authentication software using the False Acceptance Rate (FAR) and False Rejection Rate (FRR) metrics discussed in the User Authentication lecture.

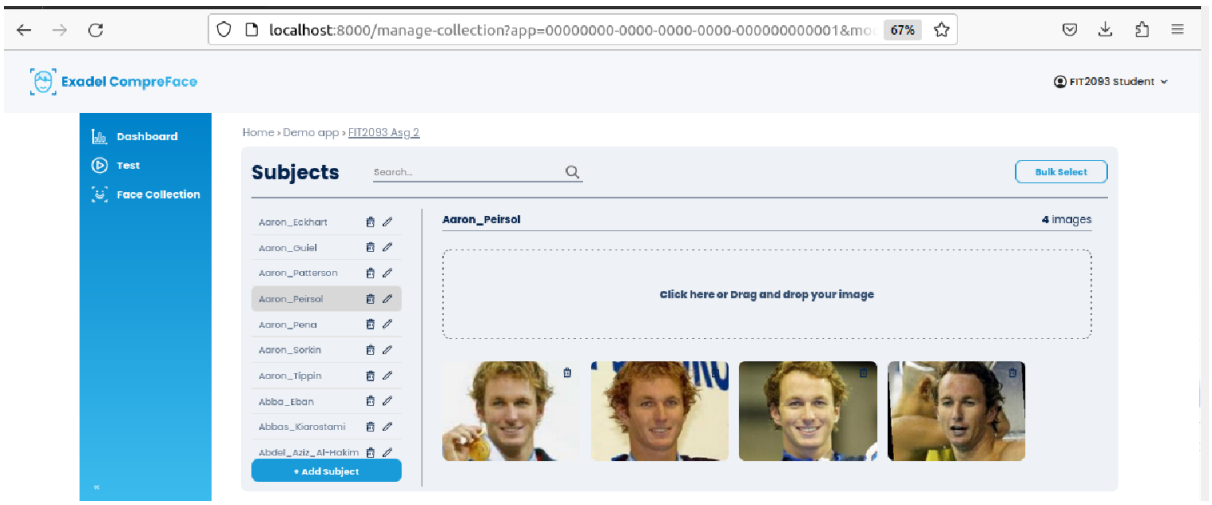
Instructions:

Option 1

1. Registering your subject:

- Install the Assignment 2 VM, and open the url http://localhost:8000/login in the Firefox browser to access the **Compreface[1] application. Use the following credentials to log in:**
  Username: fit2093@monash.edu
  Password: fit2093fit2093
- After logging in, you will see the following interface after selecting “Demo app” and then “FIT2093 Asg 2”. Over 5000 subject face images (including faces of celebrities) have been uploaded in this app[2].
- In the left panel, select “Face Collection”, to show a number of subjects (which are now the names of celebrities) pre-uploaded in the app. After clicking “Add subject”, type in your name as a subject and upload 5 of your own images under this subject.

[1] https://exadel.com/solutions/compreface
[2] Source of image: https://www.kaggle.com/datasets/jessicali9530/lfw-datase



2. Testing Phase and Computation:

- You have to prepare another set of your 10 own images and another 10 images of any other person(s) who is not in the “Face Collection”.
- Select “Test” on the left panel, then upload the testing images with the example below using a threshold score value of **0.95**. For each testing image, the CompreFace system will compare it with every registered subject image and display the best-fit registered image with the highest similarity to the testing image, i.e. the registered image with the highest similarity probability score (a number between 0 and 1). If the similarity probability score of the best-fit registered image is higher than the threshold score value, the system will consider the testing image to be a match with the best-fit registered image subject (“accept”). Otherwise, the system will regard the testing image to not match any registered subject (“reject”).

  In the example shown in Figure 1 below, the similarity probability score 0.97 of the testing image being “Adulallatif” is computed by the CompeFace app and displayed, since “Adulallatif” is the best-fit registered image to the testing image. As the similarity probability score of 0.97 is larger than the pre-set acceptance threshold score of 0.95, the testing image is falsely accepted to be “Adulallatif” although the testing image is actually of a different subject (Brad Pitt).

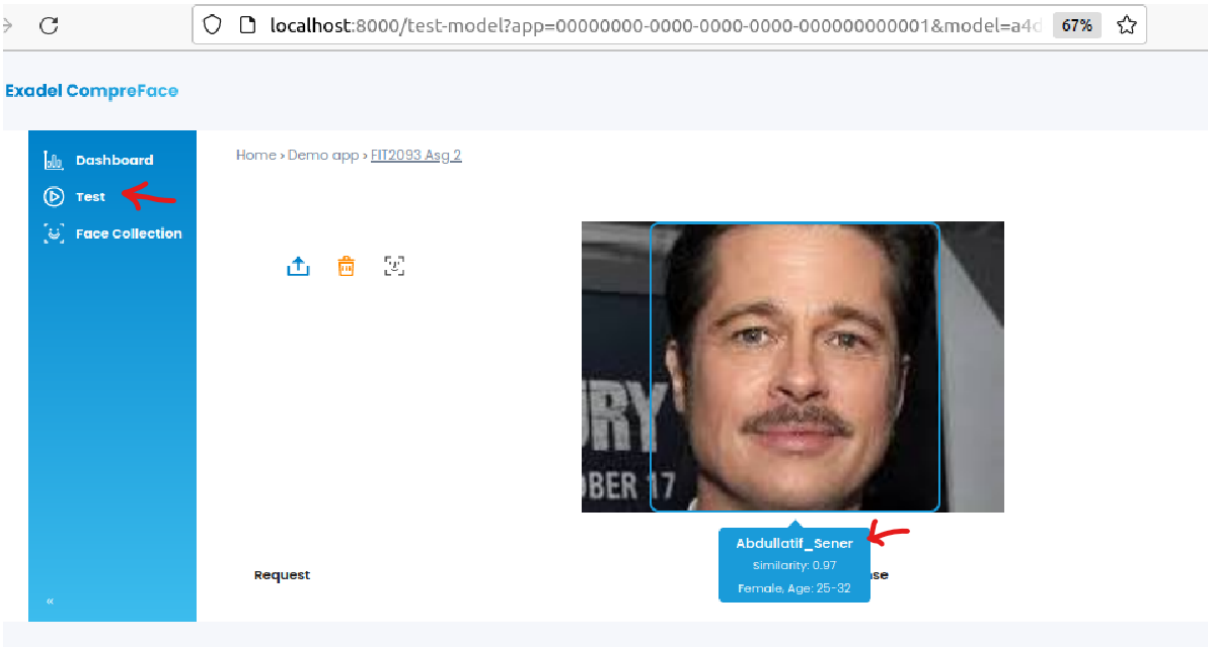


Figure 1. Example testing phase computation with CompreFace system.

- Repeat this process for all your testing images and count the number of images falling in each of the four categories below. In the above example, the FP category count will increase by 1, since 0.97 > 0.95 and the image of Brad Pitt is mistakenly recognized by Compreface as Adulallatif.

    - False Accept/Positive (FP): Probability score > threshold (“accept”) AND test image subject does NOT equal best-fit registered image subject
    - True Reject/Negative (TN): Probability score < threshold (“reject”) AND test image subject does NOT equal best-fit registered image subject

    - True Accept/Positive (TP): Probability score > threshold (“accept”) AND test image subject equals best-fit registered image subject,
    - False Reject/Negative (FN): Probability < threshold (“reject”) AND test image subject equals best-fit registered image subject

- Compute the FAR and FRR metrics for your testing images using threshold of 0.95 **and** 0.98.

- Submit a video (with your facial inset) of up to 5min to illustrate your testing process and present your results of FRR and FAR. Discuss and explain the significance of the threshold, and which value of the threshold you think would be preferred, explaining your reasons.

- In the Task 1a section of your report, include the screen captures of your images uploaded to the face collection, a table of the testing results and your calculations of

FAR and FRR with these two different threshold settings. Finally, explanation of the significance of the choice of threshold and the impacts of the choice of threshold on the security and usability of the authentication system.

Option 2

Consider the sample testing results obtained using a Face recognition software in Table 1 and Table 2.

Table 1 contains the face recognition testing results using 10 images of a **registered** person called “Alice”. For each of those 10 testing images, the table shows the name of the closest matching registered user name and the corresponding similarity probability score between the “Alice” testing image and the image of the closest matching registered user.

Table 2 contains the face recognition testing results using 10 images of an **un-registered** person called “Charlotte”. For each of those 10 testing images, the table shows the name of the closest matching registered user name and the corresponding similarity probability score between the “Charlotte” testing image and the image of the closest matching registered user.

Registered Person (“Alice”) Testing Image ID	Similarity Probability Score	Closest Matching Registered User Name
1	0.996	Alice
2	0.997	Alice
3	0.984	Candice
4	0.977	Delta
5	0.996	Alice
6	0.999	Alice
7	0.982	Eve
8	0.986	Alice
9	0.990	Alice
10	0.995	Alice

Table 1. Results for Registered Person (“Alice”) Testing Images

Un-Registered Person ("Charlotte") Testing Image ID	Similarity Probability Score	Closest Matching Registered User Name
1	0.952	Alice
2	0.937	Candice
3	0.931	Eve
4	0.918	April
5	0.915	June
6	0.937	Sara
7	0.926	Delta
8	0.909	Bella
9	0.982	Samantha
10	0.943	Samantha

Table 2. Results for Unregistered Person ("Charlotte") Testing Images

- For the testing data in Table 1 and Table 2, count the number of images falling in each of the four categories below.
    - **False Accept/Positive (FP):** Probability score > threshold ("accept") AND test image subject does NOT equal best-fit registered image subject
    - **True Reject/Negative (TN):** Probability score < threshold ("reject") AND test image subject does NOT equal best-fit registered image subject
    - **True Accept/Positive (TP):** Probability score > threshold ("accept") AND test image subject equals best-fit registered image subject,
    - **False Reject/Negative (FN):** Probability < threshold ("reject") AND test image subject equals best-fit registered image subject
- Based on the results in Table 1 and Table 2, Compute the FAR and FRR metrics for your testing images using threshold of 0.95 **and** 0.98.
- Submit a video (with your facial inset) of up to 5min to explain how you computed the FAR and FRR and present your results of FRR and FAR. Discuss and explain the significance of the threshold, and which value of the threshold you think would be preferred, explaining your reasons.
- In the Task 1a section of your report, include and explain your calculations of FAR and FRR with the two different threshold settings of 0.95 and 0.98. Finally, include an explanation of the significance of the choice of threshold and the impacts of the choice of threshold on the security and usability of the authentication system.

Task 1b) (5% of unit marks): Password Authentication

In this task, you will attempt to hack some system passwords using John the Ripper (command **`john`**) which is pre-installed in your VM (either the original Asg2 VM or the FIT2093_Asg2_LW VM; those two VMs contain identical files for Task 1b) with a built-in password dictionary, and investigate how this time and the time to compute a single password hash depends on the hashing parameters. You can time the brute force password search process by using **`time john <password file>`**. You can also measure the time taken to compute a single password hash by **`SHA-512`** using command **`time mkpasswd -m sha-512 <password>`**. Note that you should use "user time" (time spent by CPU to execute the mkpasswd process, excluding kernel and other processes' CPU time) in your below computations.

- There are **four** hashed passwords, which are hashed[3] by command `mkpasswd` using the SHA512 algorithm, located at /home/fit2093/Asg2_Task1b. For each of those password hash files, try to use the John the Ripper tool to time how long it takes to find the password by a search through the built in password dictionary of John:
  - no_salting.hash : No salt with default no. of rounds which is 5000
  - salting.hash : With salt with default no. of rounds of 5000
  - salt_1000.hash : With salt with 1000 rounds (this is the minimum allowed number of rounds)
  - salt_50000.hash : With salt with 50000 rounds

- Create and measure the time for a single password hashing using command **`time mkpasswd -m sha-512 <password>`** in the VM:
  - no salt with default no. of rounds which is 5000
  - With salt with default no. of round
  - With salt with your preferred no. of rounds
  - With salt with your higher no. of rounds

- In the Task 1b section of your **report**, include
  (1) your results of the above four types with the screen captures of the time used by John the Ripper to find the password using a brute force search through its dictionary, as well as your screen capture of the measured "system" time taken to compute a single hash (which is an estimate for the time needed by a server for a single password login verification) for each of the four types of hashing approaches using mkpasswd,
  (2) your discussion and comparison of the differences among all four approaches, you may use `man mkpasswd` to study how to use salting and change the number of the rounds; include an estimate of the time for a brute force search through a dictionary of 200 Million passwords for each approach
  (3) your recommendation of which approach should be used for password hashing and your reasoning based on usability and security considerations

---

[3] See this page for more information on the the crypt function used in the mkpasswd command of Ubuntu Linux for password hashing, including the default number of rounds: https://manpages.ubuntu.com/manpages/bionic/man3/crypt.3.html

- In the Task 1b section of your **video** of up to 5 min, include
  (1) a summary of all results of the time used by John the Ripper to find the password using a brute force search through its dictionary, as well as the measured time taken to compute a single hash (which is an estimate for the time needed by a server for a single password login verification) for each of the four types of hashing approaches using `mkpasswd`,
  (2) discussion and comparison of the differences among all four approaches and the time estimate for a brute force search (without salting, or once the salt is exposed in the hash table) through a dictionary of 200 Million passwords with the different choices of number of rounds,
  (3) your recommendation of which approach and how many rounds should be used for password hashing and your reasoning based on usability and security considerations.

Task 2: (10% of unit marks): Access Control

In this task, you will assume the role of a system administrator, and your task is to configure the permissions for two new users and test if the file permissions are set appropriately for the desired access control policy. You have to create a video (with your facial inset) of no more than 5 minutes to demonstrate your answers to Task 2 (a) to (c). In your video presentation, you may use Powerpoint or any other software to explain your answers to the tasks if needed. Your report should include the screen captures, explanations and the command used in all subtasks in Task 2. Note that either the original Asg2 VM or the FIT2093_Asg2_LW VM can be used in this task; those two VMs contain identical files for Task 2.

2a) (4% of unit marks) Create two new `users`, called peter and mary, who each have their home directory, called `/home/[username] (e.g. /home/mary)`. Use command `useradd` to add mary into the existing groups `hr` and `it`, and add Peter only into the group `it`.
In your video, you should:
(1) Show your steps of creating the two users and adding them into the corresponding groups.
(2) Show the contents of the file `/etc/group`.
(3) Modify `hr.txt` using text editor gedit or nano or vi in the folder `/home/share-folder/hr` and create the file `it.txt` in folder `/home/share-folder/it` as the user mary.
(4) Modify `/home/share-folder/hr/hr.txt` and `/home/share-folder/it/it.txt` as the user peter.
(Note: to remove a user, you can use the command "`userdel -r [username]`")

2b) (3% of unit) A program, `readsecret`, is to read the file `secret.txt` in folder `common`. Login as `fit2093` to set UID of the program such that `mary` and `peter` can run it to acquire the secret in "`secret.txt`".

In your video, login as the two users (peter and mary) and run the program, `readsecret`, owned by `fit2093` in `common`. Show the file's permissions and explain why the two users can extract the secret in `secret.txt`.

2c) (3% of unit) Folder `employee` allows all user groups to read the file, `readonly.txt`, however, only users in group hr can modify it. In this part of the task, you should login as user `peter` to modify the file `readonly.txt` without the change of permission settings. Suggest a mitigation if peter can modify the file.

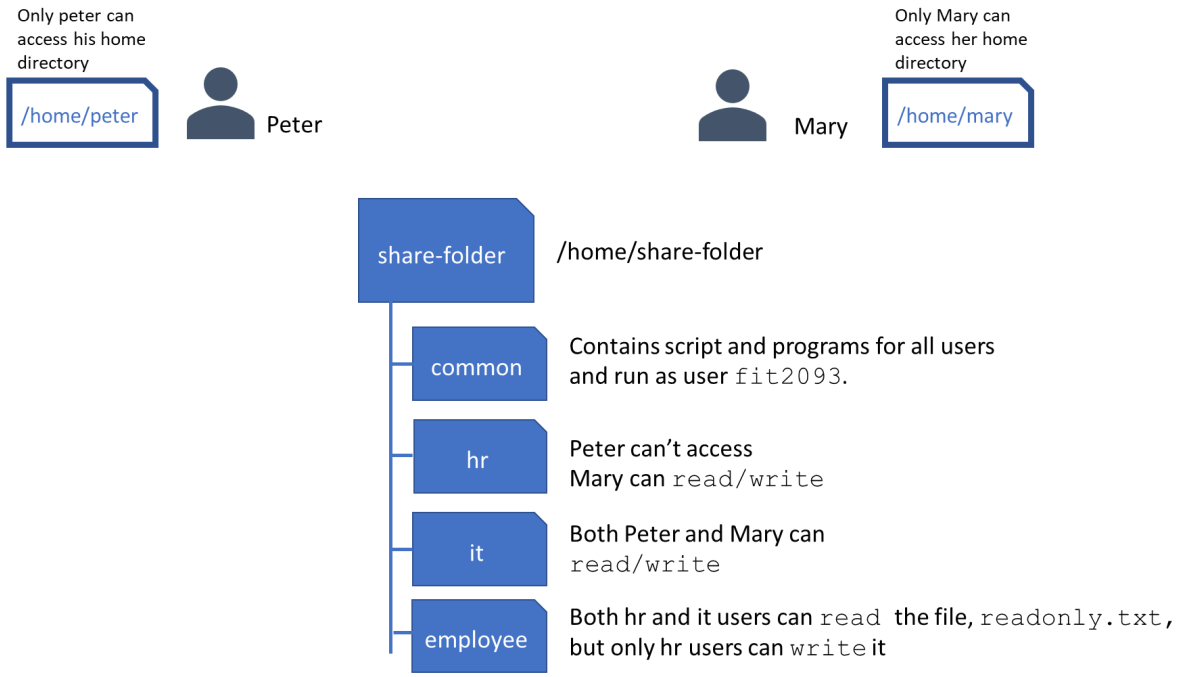


Figure 2. The desired access control requirements for the file system.

WARNING (Academic integrity): It is an academic requirement that your submitted work be original. Zero marks will be awarded for the whole submission if there is any evidence of plagiarism or contract cheating (i.e. paying another person to complete the assessment task). It is fine to use code or other material from various sources in your report. However, any material that you obtain from some source (e.g. website, book, paper, article) **must be cited** in the appropriate place in your report **and listed in the references section** of your report. Please also note that students must not work on this assignment with members of other assignment groups, and significant similarities between assignments submitted by different groups (other than those due to the same cited starting source code / dataset) will be investigated for evidence of plagiarism.

REMARK (Guidelines on Use of AI tools in the Assignment): ChatGPT or other AI tools may be used for study purposes, to learn about your topic, and to develop your assignment. However, similar to citation requirements for other references (see “Academic Integrity” statement above), you **must include a clear declaration of all generative AI tools used** (e.g. ChatGPT, DALL-E, Grammarly, voice-to-text), **how and where you have used them**. In particular, you should be aware that output of AI tools may not be factually correct and you should therefore critically evaluate the output generated by such tools for claim accuracy and appropriateness to the topic, using reliable sources, before incorporating such output in your assignment (e.g. an example declaration may be: ‘ChatGPT was used to generate an initial structure for the Introduction and Conclusion. I then edited this to correct factual inaccuracies, add citations to support claims, and strengthen the connection to my chosen topic and the ideas from other references that I referred to’).

Where to get help

What can you get help for?

English language skills

if you don't feel confident with your English.

- Talk to English Connect: https://www.monash.edu/english-connect

Study skills

If you feel like you just don't have enough time to do everything you need to, maybe you just need a new approach

- Talk to an academic skills advisor: https://www.monash.edu/learnhq/consultations

Things are just really scary right now

Everyone needs to talk to someone at some point in their life, no judgement here.

- Talk to a counsellor: https://www.monash.edu/health/counselling/appointments (friendly, approachable, confidential, free)

Things in the unit don't make sense

Even if you're not quite sure what to ask about, if you're not sure you won't be alone, it's always better to ask.

- Ask in the forums or email your tutor:

  Teaching team: https://lms.monash.edu/course/view.php?id=155649&section=1

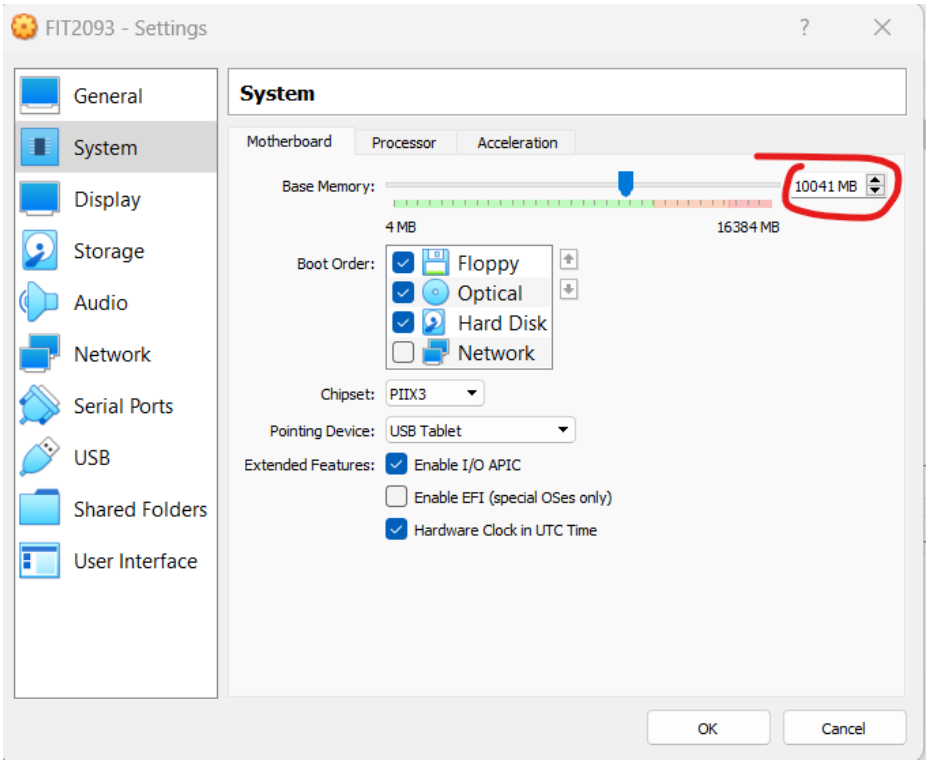
  Consultation: https://lms.monash.edu/mod/resource/view.php?id=11630825

I don't know what I need

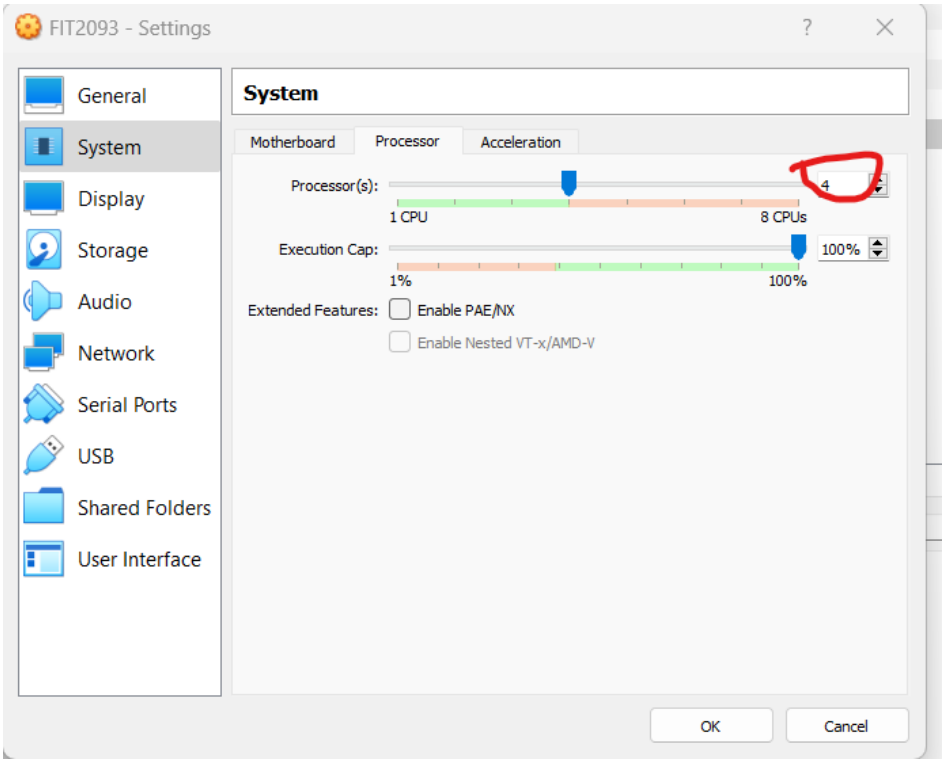
Everyone at Monash University is here to help you. If things are tough now they won't magically get better by themselves. Even if you don't exactly know, come and talk with us and we'll figure it out. We can either help you ourselves or at least point you in the right direction.

Appendix

Update the system's base memory to 9216 MB or the higher in your computer.



Set processor to 4 or the no. of processors in your computer. (recommend 4)



Change log

All changes to the assignment will be listed here with the time of the change (in Melbourne time):

- **26 April 2023**: version 1.1 of this Asg2 spec document was released, with an extended deadline and a new submission option (Option 2). Option 1 is identical to the original (version 1.0) specifications document, **except for a couple of corrections to the Task 1b, namely: using "user time" for timing (not "sys time" as stated previously), and the default number of rounds is 5000 (not 9 rounds as stated previously)**. All changes introduced in version 1.1 are shown in purple font.