

HW2 (100 POINTS)

Before you start:

1. Download the zip file HW2.zip. This file will usually be saved in the Downloads folder (directory).
2. Move HW2.zip to your home directory by using GUI-drag and drop.
3. Open the zip file. A folder named HW2 will be created in your home directory.
4. Open a terminal session.
5. Go to your HW2 directory by typing in “cd HW2”
6. Use vi to make and edit files in directory HW2.
7. Open a second terminal by pressing **CMD+n** and use this terminal to test the Unix code that you write as answers to the questions.

Follow these instructions in each file you submit

- **Include the class header**
- **Format of Q&A pairs**

Example:

```
#mprocop2:09/01/2024:filename
```

```
#Q1
```

```
cat file
```

```
#Q2
```

```
ls -lt file
```

This HW has 3 parts, HW2-part1, HW2-part2 and HW2-part3, and will have 3 separate submissions on Gradescope.

All these parts are due Sunday Feb 4th 11:59pm

If you submit:

- HW2-part1 by Wed Jan 31st (11:59pm) will get 3 extra points
- HW2-part2 by Friday Feb 2nd (11:59pm) you will get 3 extra points
- HW2-part3 by Saturday Feb 3rd (11:59pm) you will get 2 extra points

Work on HW problems on your own. Do not share your code with other students. Always double-check the scripts you submit on Gradescope.

Ubuntu users. Please include a comment line in each script stating that you use Ubuntu. Include the statement down below the class header, like this:

```
#class header
```

```
#Ubuntu
```

General Grading:

All points for correct code

All points off for syntax errors

0.5 points for class header

0.5 points for correct Q&A format

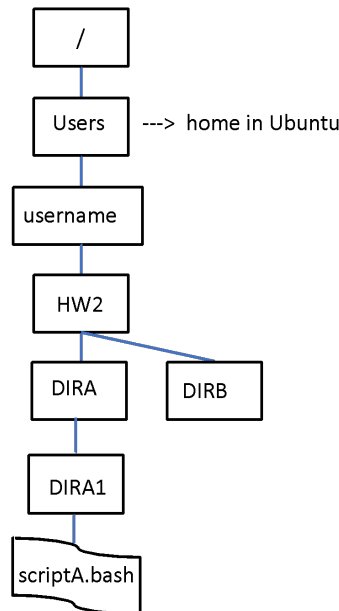
-0.5 for case sensitive error

-1 literally write your-username

We take points off if you do not use class material.

HW2-part1 (28)

1 (13) Make a file called **Ex2_1.bash** and answer the following questions. Make sure to start this exercise from your HW2 directory. Within HW2 you will find the following files and directory structure:



Keep in mind that what we call home directory is the directory named with your username. Write pathnames as they work on your computer. Answer each question with only one line of code.

Q0. (0) Go to the HW2 directory. You must start from the HW2 directory.

Q1. (3) From the HW2 directory, run the script **scriptA.bash** by using a relative pathname. Do not change your working directory.

Q2. (4) From the HW2 directory, use the find command to locate the **hidden directory** called **.myhidden**. Start the search from your working directory and use a relative path.

Q3. (2) View the content of the file fileA, which is within DIRA. Use an absolute path.

Q4. (3) List in long format the content of DIRB

2 (15) Make a bash script called **Ex2_2.bash**, and in it do the following:

Q1. (1) Define a variable called **v1** and assign it the value 10.

Q2. (2) Make another variable **v2** that is equal to the result of the arithmetic operation:
 $(v1 - 5) * 4$.

Do not use the number 10 but use variable **v1** and shell arithmetic.

Q3. (3) Use shell arithmetic, variables **v1** and **v2**, and the **echo** command to print to screen the result of this arithmetic operation $(v2 * v1) / 2$ together with the text:

The result is the number 100

Do **not** write the number 100. Instead, this number should be the result of the shell arithmetic operation stated above.

Q4. (3) Make a variable called **var** that is equal to this string:

I love

Q5. (2) Make a variable called **music** that is equal to your preferred music.

Q6. (3) Use the **echo** command to print to screen:

I love *yourmusic* music

Example: if you store *jazz* in variable **music**, the output should be

I love *jazz* music

Do not write *yourmusic* and do not write *I love* but use the variables **var** and **music** that you made above.

After you edit and save the script, run the script, and make sure that there are no error messages. In case of error messages, fix the errors.

The **only** output of the script should be:

```
The result is the number 100
I love jazz music
```

Upload the following files to Gradescope **HW2-part1**:

Ex2_1.bash

Ex2_2.bash

HW2-part2 (30)

1 (30) Make a file called **Ex2_3.bash** and in it answer the following questions. Make sure that you start this exercise from your HW2 directory. **Use relative pathnames and meta-characters to answer the questions**, and in all cases answer questions with only one line of code. Directory DIRB is within the HW2 folder.

Q0 (0) Go to HW2 directory.

Q1. (4) Your working directory is HW2. List all the files in the DIRB directory whose names start with either the letter a or letter b and have 2 characters. Do this with one line of code. Do not change your working directory. Use a relative pathname. Do not use the find command.

Q2. (4) Your working directory is HW2. List all the files in DIRB directory whose names do not start with an g. Do this with one line of code. Do not change your working directory and use a relative pathname. Do not use the find command.

Q3. (4) Your working directory is HW2. **View the contents** of all the files in the DIRB directory, whose names end with .dat. Do this with one line of code. Do not change your working directory and use a relative pathname. Use a file viewing command. DO NOT USE ls.

Q4. (4) Your working directory is HW2. Remove all files in the DIRB directory whose names have only 3 characters. Use an option of the rm command to enable confirmation of deletion. Do this with one line of code. Do not change your working directory and use a relative pathname. Do not use the find command.

Q5. (5) Your working directory is HW2. Copy all files that are in the DIRB directory, whose names start with s and have only 4 characters, into directory DIRA. Do this with one line of code. Do not change your working directory and use a relative pathname.

Q6. (4) Your working directory is HW2. Create 10 directories inside directory DIRA named dir1, dir2, ..., dir10. Do this with one line of code. Do not change your working directory and use brace expansion and a relative path. Use metacharacters properly.

Q7. (4) Your working directory is HW2. Make files fileB, file10 and fileA1 only within each of the following directories dir1, dir5 and dir7. Use brace expansion to specify directories and names of the files. Do this with one line of code. Do not change your working directory and use brace expansion and a relative path. Use metacharacters properly.

Upload the following files to Gradescope **HW2-part2:**
Ex2_3.bash

HW2-part3 (42)

1. (18) The provided data file [6KX7.pdb](#) was downloaded from the Protein Data Bank (PDB).

The PDB is a database for the three-dimensional structural data of large biological molecules, such as proteins. The data is typically obtained by X-ray crystallography and NMR spectroscopy and then submitted by biologists and biochemists from around the world. The PDB is a key in different areas of structural biology.

The data file is in pdb format (a textual file format) describing protein and amino acids, including atomic coordinates, secondary structure, as well as atomic connectivity.

The data file 6KX7.pdb contains information on the mouse Cryptochrome 1 (CRY1), a circadian clock protein, which is an essential component of the circadian clock controlling daily physiological rhythms.

Amino acids are the structural units (monomers) that make up proteins and are identified by different codes. In this exercise, we will be using a three-letter Letter Code.

Let's extract information from this data file by using bash text processing commands. If you want general info on the pdb format, you can click [here](#).

Make a file called **Ex2_4.bash** and answer the following questions. The data file 6KX7.pdb is within HW2. **Use only one line of code.**

Q0. (0) Change to HW2 directory.

List the content of the HW2 and you will find the data file 6KX7.pdb

Q1. (4) Use options of the wc command to display both the number of lines and the number of bytes of the file 6KX7.pdb.

Q2. (6) Use grep to extract all the lines containing keyword EXPDTA or keyword TITLE or keyword HEADER

Q3. (5) Use grep to count the lines that do not contain keyword SEQRES.

Q4. (2) Use grep to extract all the lines containing the amino acid code (keyword) GLU

2. (24) Make a file called **Ex2_5.bash** and answer the following questions. You need the file `amino_hydro_weight.dat`, `file1` and `file3` which are within HW2.
The provided data file **amino_hydro_weight.dat** contains information about the hydrophobicity scale (of Kyte and Doolittle) and weight (in g/mol) of the amino acids.

The file is organized as follow:

1st field: amino acid one-letter code
2nd field: hydrophobicity
3rd field: weight

Look at the content of the file before starting this text processing exercise.

Q0. Go to HW2 directory and answer all the questions from HW2.

Q1. (6) Sort the file **amino_hydro_weight.dat** based on decreasing weight (highest weight on top).

Q2. (5) Use `cut` to extract both the 1st and 3rd fields of the file **amino_hydro_weight.dat**

Q3. (3) Use the provided files `file1` and `file2` and a Unix command to obtain the following output to screen. You should use brace expansion.

```
Brazil*Chile*Brazil
```

Q4. (2) Use files `file1` and `file2` and a Unix command (do not use `paste`) to obtain:

```
Brazil
Chile
Brazil
```

Q5. (3) Use `sed` on `file2` to display to screen:

```
Vinile
```

Q6. (4) Make these variables:

```
price=173.531
stock=IBMComp
date=5/6/2011
```

After you make the variables, use `printf` in `bash` and the variables to print to screen this formatted text:

```
highest price=$173.5 stock=IBM date 5/6/2011
```

You must define the variable in your solution script.

You should have 4 lines of code for this part. You can format the date as a string.

Upload the following files to Gradescope **HW2-part3**:

Ex2_4.bash

Ex2_5.bash