

## Zhenhan Zhao

---

**General Instruction:** To complete the homework set, you are required to do the followings. Your solutions must be typed in L<sup>A</sup>T<sub>E</sub>X using the course homework template. The progression of your homework solution is to be “recorded” by making a git folder specifically for this homework set. The burden of proof is on you, and if your git commit history is sparse, then you may be liable for a penalty. A paper copy of the PDF output of your L<sup>A</sup>T<sub>E</sub>X file is to be submitted to your instructor in class on the due date. *After* submitting the paper copy, but *before* the end of the due date, you will upload your work to your github by making a remote repository specifically for the homework, and post the link to the repository at the designated *Discussion* forum in Blackboard by making a thread just for you. The repository name in your github should be `550400.homeworkset.1` and the discussion forum thread should be named `YourFirstNameMiddleInitialLastName`, e.g., `BaracHObama` and `WillardMRomney`. You have till the end of the due date to finalize your github repository. However, any commit made after the class time of the due date will be inadmissible. *Your attention to details in following this instruction will be critical, and if not followed exactly at the time of collection, the homework set may be graded at 90% of the full score.*

**Problem 1 (10 pts):** Assume that you are starting from “scratch” at the directory `~/`. Provide a sequence of git/bash commands that yields a git folder with a commit history such that:

- the *master* branch has commits *A*, *B*, *C*, *X* and *D*,
- the *alt* branch has commits *A*, *B*, *X*,

Suppose that you are currently working on `master` branch. Draw its commit history graph (i.e., the graph portion of the output of `git log --graph --oneline`). Next, assume that you are on `alt` branch. Draw its commit history graph. **Answer:**

```
cd 550400
mkdir hw1
cd hw1
git init
vi main.txt
git add .
git commit -m "A is done"
vi main.txt
git add .
git commit -m "B is done"
git branch alt
git checkout alt
vi main.txt
git add .
git commit -m "X is done"
git checkout master
```

```

vi main.txt
git add .
git commit -m "C is done"
git merge master alt
vi main.txt
git add .
git commit -m "merge is done"
vi main.txt
git add .
git commit -m "D is done"
cat main.txt
git checkout alt
cat main.txt
git checkout master
git push https://github.com/zhaozhenhan/Homework1Problem1 master
git push https://github.com/zhaozhenhan/Homework1Problem1 alt
git log --graph --oneline

```

```

$ git log --graph --oneline
* 50dd1af D is done
*      8ee6838 merge is done
| \
|  * 9979a84 X is done
|  * | bbc5c37 C is done
|  !/
* 80be422 B is done
* f44a3da A is done

```

Figure 1: master graph

```

$ git log --graph --oneline
* 9979a84 X is done
* 80be422 B is done
* f44a3da A is done

```

Figure 2: master graph

**Problem 2 (10 pts):** Assume that you are starting from “scratch” at the directory `~/`. Provide a sequence of git/bash commands that yields a git folder and

- configure your git with your name and your email address,
- set up an alias for each of the git remotes listed below:

```

git://github.com/nhlee/550400.stanza1.git
git://github.com/nhlee/550400.stanza2.git
git://github.com/nhlee/550400.stanza3.git

```

Assume that each remote contains exactly single commit with a txt file for a single (different) stanza,

- pull to combine three stanzas of a poem,
- after the first pull, add the title of the poem,
- after the second and third pull, resolve the merge conflict,
- after resolving the third pull merge conflict, push the result to your (newly created) remote repository.

**Answer:**

```
cd ..
mkdir hw2
cd hw2
git init
git config --global user.name "zhaozhenhan"
git config --global user.email "zhaozhenhan@gmail.com"
u vi main.txt
git add .
git commit -m "alias is done"
git pull https://github.com/nhlee/500400.stanza1
vi main.txt
git add .
git commit -m "s1 and title is done"
git pull https://github.com/nhlee/500400.stanza2
vi main.txt
git add .
git commit -m "s2"
git pull https://github.com/nhlee/500400.stanza3
vi main.txt
git add .
git commit -m "s3"
git push https://github.com/zhaozhenhan/Homework1Problem2 master
```

**Problem 3 (40 pts):** Consider a team of four students, say,  $A$ ,  $B$ ,  $C$  and  $D$ , who just started working on writing a `latex/beamer` file, say `main.tex`, for a class presentation of their work statement. Assume that they do not wish to coordinate their schedules for a concurrent group meeting (both virtually and physically). Assume that:

- $A$  is in charge of *Introduction*,
- $B$  is of *Problem Statement*,
- $C$  is of *Timeline*,
- $D$  is of *Deliverable* part of the presentation.

In other words, their contributions to `main.tex` do not overlap. Then,

- first, devise a work flow strategy for the team so that they can collaborate asynchronously using `git`,
- next, devise yet another `git` strategy different from your earlier proposal.

Finally,

- discuss the strength and weakness of each of your proposed strategies in terms of merge conflicts resolution,
- make the final recommendation.

In order to answer this question, *build* a mathematical model, *following* the guideline from IMM. Use Section 1.4 and Section 1.5 of IMM as *role models*. For example, you are to identify which variables are exogenous and which are endogenous. More specifically, among other things, in your model, is the preamble part of `main.tex` an endogenous or exogenous variable? Note also that in addition to this issue, there are other issues that you are to consider. So, *be sure to consult IMM*.

**Answer:**

One strategy is to establish four branches representing A,B,C,D four students after we initiate a git folder. For each one of the students, after they finish their work, they have to pull from the master first, then merge his work with others and then push back to master. Following is the specified figure,

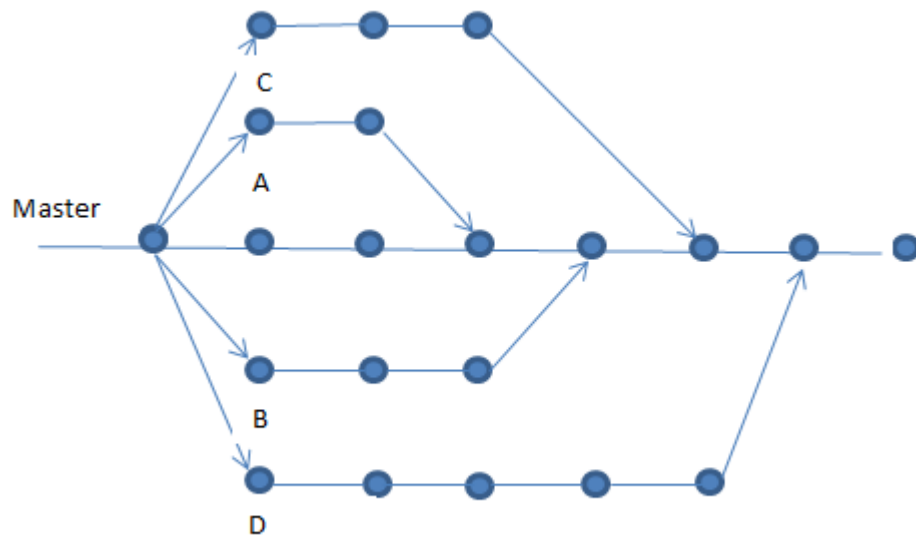


Figure 3: stragey 1

Next, for another strategy, we build four git folders and A,B,C,D can work at branch master under their own folder. After they finish their work, a random person from the four will pull the others's work and merge them together to finish the report, the following is the specified figure,

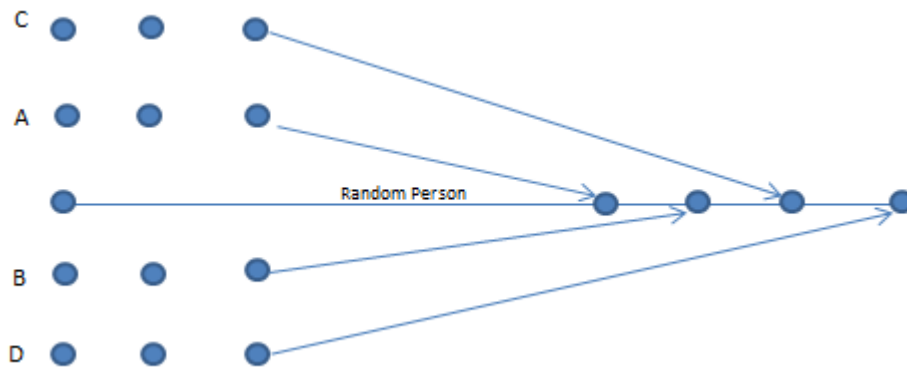


Figure 4: stragey 2

In this problem, what we want to know is what kind of strategy is better, in terms of the effects of merging.

The important parts of the model is the merging of four students' work, except this, other parts of the main.txt are not important and are treated as exogenous variables.

For both strategies, they both could generate a complete report. As for strategy 1, everyone works independently and they all will pull, merge and push to finish the report, everybody's work will be included in the report. As for strategy 2, we will pick a random student to do the pull, merge and push once everyone finishes their part. But how to randomly pick a student? For me, I prefer the first strategy, everybody gets to practice the pull, merge and push so their work burden is equal in a qualitatively way.

To test the model, we can assign simple sentences to each work and follow the work flow of each strategy to analyze which one works more efficient in merging.

**Problem 4 (aka. Fair Play, 40 pts):** Answer the following question:

Is the tennis game fair?

Note that unlike Problem 3, this question is vaguely stated. This is intensional, whence to begin, you will first need to clarify what exactly your question is. You may use the class discussion on this particular problem, but you *may not* directly refer to our discussion. Instead, formulate the model carefully but concisely in your own words.

**Final Remarks about Problem 3 & Problem 4:** They are open-ended problems. However, your scores will be determined by how well do you follow the exposition style outlined by IMM and WMA. For both problems, your write-up should be

- self-contained,
- covering all four parts of Section 1.3 of IMM,
- paying a particular attention to any causal relation that you might be investigating, following Chapter 3 of WMA,
- answering questions that are explicitly asked in the problem statements.

For Problem 3, focus mostly on Step 2 and Step 3 of Section 1.3 of IMM. For Problem 4, focus mostly on Step 1 and Step 2. For each problem, minimum 1 pages and maximum 2 pages.

**Answer:**

1. Formulate the problem. Before formulating the problem, we need to clarify some ground rules for the tennis games. In the tennis game, we have a tennis match which takes place between the server who serves the ball to start the point called server and receiver who receives the ball served. It is always the server who wins the first score, which is 15, if he wins the second score, he gets 30, similarly for the third and fourth score, he gets 40 and 45. If we have a tie at 40-40, the winner must win by two scores. So the problem we are facing is is the tennis game fair? First, let's rule out some exogenous factors. Say, one player is more skillful or the other player has injured before or the other has never played in clay court and gets pretty nervous about it, then we have one side who stands a better chance to win and of course the game isn't fair. But in our problem, we rule out these factors including the skills, injury, previous experience and psychological state. We are only interested in whether the game is fair if we assume these factors are the same for both players.

2. Outlien the model. First, we pick up two players with the same tennis skills, health and psychological conditions presumably and then have them play  $N$  games. For each game, we assign value 1 to the server if he wins and assign value 0 if he loses. To test whether the game is fair equals to the question whether the probability of winning for the server is 0.5. After  $N$  games, we sum up all the values and then divided it by  $N$  to check if it's close to 0.5. If it is, we conclude that the game is fair. To give more accuracy to the model, we could have two players play more games, as long as  $N$  is larger, the conclusion is more accurate. If we make  $N$  large enough and the probability isn't close to 0.5, we conclude the game is not fair.
3. Is it usefull? In the real world, the assumption we made about the skills differences and health differences will not stand and the game result may be largely influenced by these factors. However, our model is still advisable in sports statistics to assess the probability of winning for server.
4. Test the model. We could pick two player with almost the same conditions and collect the game data to test the model.