

Git Workshop

Part 2 of 3

Merge Conflicts
and
Remote Repositories

Workshop Agenda

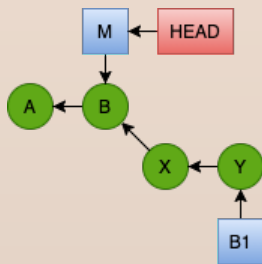
- Part 1 – Another Deck
 - Installation and Setup
 - Concepts
 - Repository Initialization
 - Clone
 - Basic Lifecycle
 - Logs
 - Introduction to Branches
- **Part 2 – This Deck**
 - Merges
 - Remote Repositories
 - Fetch and Push
- Part 3 – Another Deck
 - Tags
 - Hosting Services
 - GitHub
 - AWS
 - Commit Signing
 - Author Delegation
 - Stashing

This is a "presentation-ification" of the HTML workshop available at
<https://www.workshops.lacounty-isab.org/>

Scenario Continued

This Part 2 continues a scenario that was started in Part 1. If you still have your files and folder from Part 1, you can skip to the next slide now. This slide explains how to recreate where we left-off from Part 1.

Part 1 ended with a merge of the situation shown on the right. It resulted in a *fast-forward* merge where the **master** branch pointer advanced to the **Y** node.



The first exercise of this Part 2 will explore an alternative to the Part 1 ending – the consequence of adding commits to the **master** branch *before* the merge. We need the branch pointers reset to how they were before the merge updated them. Step 8 of Exercise 6 cloned the state of the repository to a `samples2` directory before proceeding. If you no longer have that `samples2` directory available, the steps on the right will show you how to recreate it from a repository hosted on GitHub.

1. Change to your `GitWorkshop` directory. This is a directory in which we undertake these workshop exercises.
2. If you have a `samples2` subdirectory, you may not need to follow these steps. It may have been created during your Part 1 session. If you're unsure, delete the `samples2` directory.

3. Clone the **lacounty-isab/gitwkspex6** repository to `samples2`.

```
GitWorkshop$ git clone
https://github.com/lacounty-isab/gitwkspex6.git samples2
```

4. Change to `samples2`.
5. List branches. You'll find that **master** is the only local branch. Create a *local* **B1** branch from the **origin/B1** remote branch*.

```
GitWorkshop/samples2$ git branch -av
* master                d0d9eea Added Python and fixed typos.
remotes/origin/B1       dbc1acc Songs to compositions.
remotes/origin/HEAD     -> origin/master
remotes/origin/master   d0d9eea Added Python and fixed typos.
GitWorkshop/samples2$ git branch B1 origin/B1
Branch 'B1' set up to track remote branch 'B1' from 'origin'.
GitWorkshop/samples2$ git branch -av
B1                      dbc1acc Songs to compositions.
* master                d0d9eea Added Python and fixed typos.
remotes/origin/B1       dbc1acc Songs to compositions.
remotes/origin/HEAD     -> origin/master
remotes/origin/master   d0d9eea Added Python and fixed typos.
```

* *Remote branches* will be covered as part of *remote repositories*.

Turn Back the Clock

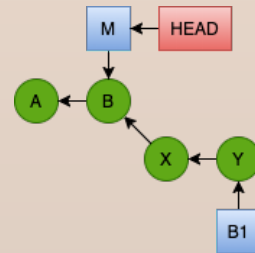
Note: If you followed the instructions from the last slide, you may skip this slide. You already have the final result.

Next we'll simulate a case where a commit was added to the **master** branch **before** our merge operation. To this end, let's go "back in time" by changing to our GitWorkshop/samples2 directory. We created this directory in Step 8 of the last exercise.

```
GitWorkshop/samples1$ cd ../samples2
GitWorkshop/samples2$ git log --oneline
81d60de (HEAD -> B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
10f629d (origin/master, origin/HEAD, master) Added Python and fixed typos.
b83eb9b Initial version.
GitWorkshop/samples2$ git branch
* B1
  master
GitWorkshop/samples2$ git checkout master
Switched to branch 'master'
Your branch is up to date with 'origin/master'.
GitWorkshop/samples2$ git branch
  B1
* master
GitWorkshop/samples2$
```

In GitWorkshop/samples2, **B1** is still the current branch (**HEAD** points to it). We also confirmed this with the **branch** command. We changed the current branch to **master** using the **checkout** command.

After all this, we're back-in-time to before the **B1** merge.



```
GitWorkshop/samples2$ git log --oneline
10f629d (HEAD -> master, origin/master, origin/HEAD) Added Python and fixed typos.
b83eb9b Initial version.
```

Add a Commit to Master

Edit `hg17.txt`. Note that the "songs" → "compositions" change on line 17 is no longer there. Edit line 10 to change

thirty-seven → forty-two

and save the file. Run the **git diff** command to confirm your change.

```
GitWorkshop/samples2$ git diff
diff --git a/hg17.txt b/hg17.txt
-bunkers some thirty-seven miles from the stage,
+bunkers some forty-two miles from the stage,
```

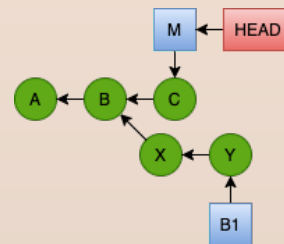
Add this change as a commit.

```
GitWorkshop/samples2$ git add .
GitWorkshop/samples2$ git status
On branch master
Your branch is up to date with 'origin/master'.

Changes to be committed:
  (use "git restore --staged <file>..." to unstage)
    modified:   hg17.txt

GitWorkshop/samples2$ git commit -m "Forty-two miles."
[master 4878aee] Forty-two miles.
1 file changed, 1 insertion(+), 1 deletion(-)
```

This latest change is represented as commit **C** in the updated diagram. When we list the log, we only see entries *reachable* from commit **C**.



```
GitWorkshop/samples2$ git log --oneline
4878aee (HEAD -> master) Forty-two miles.
10f629d (origin/master, origin/HEAD) Added Python
and fixed typos.
b83eb9b Initial version.
```

This concept of *reachable* is important for understanding many Git operations. A node in the Git graph (**D**irected **A**cyclic **G**raph or DAG) is *reachable* from a point C if it may be reached from C by traversing the edges of the graph in the direction of the arrows.

In the diagram above, commits **A**, **B** and **C** are reachable from **C**. **X** and **Y** are not reachable from **C**.

By default, the **log** command only displays commits reachable from **HEAD**.

Displaying Branch Commits

By providing a branch name to the log command, we can see all log entries for commits reachable from that branch. To see the commits reachable from **B1**:

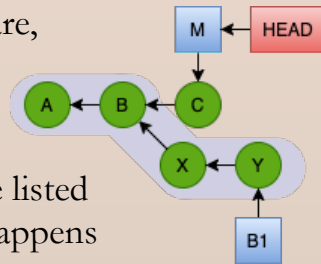
```
GitWorkshop/samples2$ git log --oneline B1
81d60de (B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
10f629d (origin/master, origin/HEAD) Added Python
and fixed typos.
b83eb9b Initial version.
```

The commits reachable from **B1** are, following the arrows:

$Y \rightarrow X \rightarrow B \rightarrow A.$

This is the order in which they are listed by the **log** command. This also happens to be reverse chronological order.

In this way, we can view logs from all the commits on any particular branch. We are not restricted to viewing the log history of **HEAD**. **HEAD** is simply the default.



But **all the commits** reachable by any particular node is usually much more than we want. Our sample repository is small. Real life repositories have hundreds of commits. We saw earlier a common way to restrict the output is with **-N** where N is a number.

```
GitWorkshop/samples2$ git log --oneline -2 B1
81d60de (B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
```

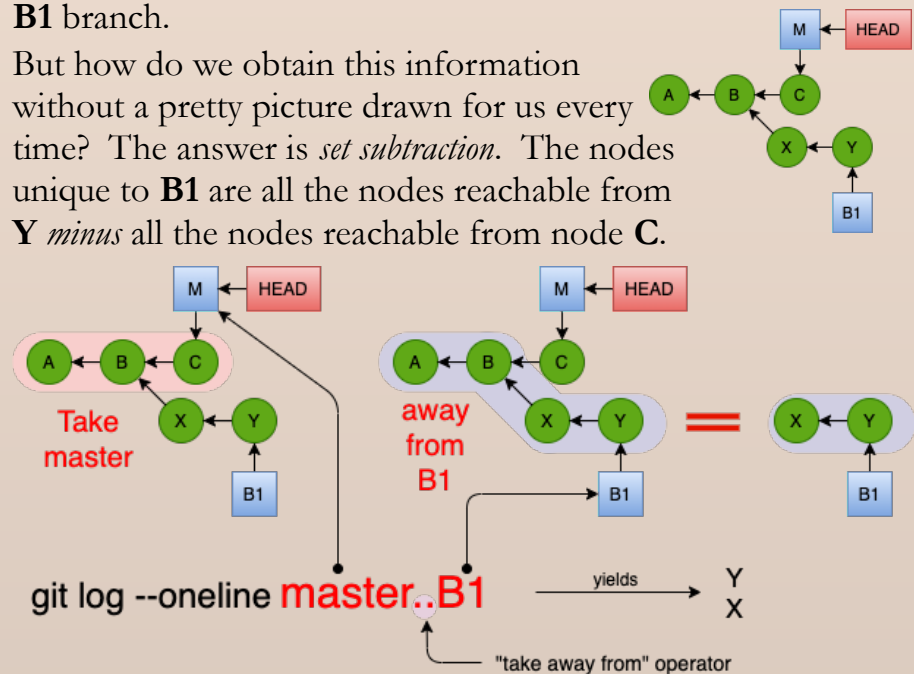
In the example above, we used **-2** to restrict the history to two entries.

Another common inquiry is: "List the commits on a branch since the branch was created." For this kind of inquiry, certain Git command support *set subtraction*, denoted by the **..** operator (two periods).

Log Set Subtraction – Two Dots

A common inquiry in a branch scenario is to list the commits that have occurred since the branch. In our current example, the branch point is node **B**. Node **C** has been added to the **master** branch and nodes **X** and **Y** have been added to the **B1** branch.

But how do we obtain this information without a pretty picture drawn for us every time? The answer is *set subtraction*. The nodes unique to **B1** are all the nodes reachable from **Y** *minus* all the nodes reachable from node **C**.



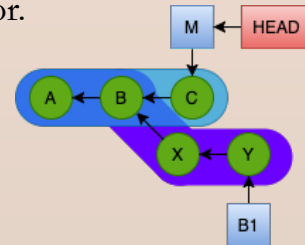
```
GitWorkshop/samples2$ git log --oneline master..B1
81d60de (B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
```

We can ask the question in reverse: which commits have been made to master since the **B1** branch? We just flip the order of the arguments in the difference operator.

Take all
nodes
reachable
from **Y**

..

away from
all nodes
reachable
from **C**



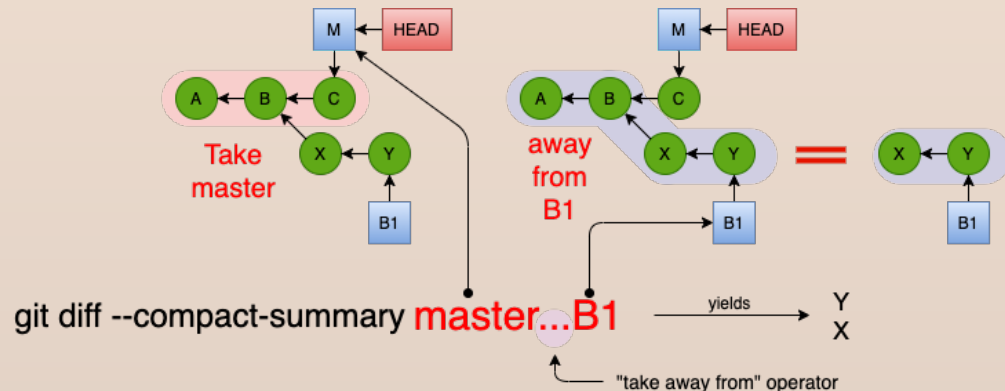
```
GitWorkshop/samples2$ git log --oneline B1..master
4878aee (HEAD -> master) Forty-two miles.
```

The difference operator assumes **HEAD** when the commit is omitted. Since **HEAD** is currently **master**, the last two commands can be abbreviated as follows.

```
GitWorkshop/samples2$ git log --oneline ..B1
81d60de (B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
GitWorkshop/samples2$ git log --oneline B1..
4878aee (HEAD -> master) Forty-two miles.
```

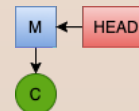
Diff Set Subtraction – Three Dots

A similar concept applies to the **diff** command, except the set subtraction operator is three dots instead of two.



```
GitWorkshop/samples2$ git diff --compact-summary master...B1
hg17.txt | 2 +-
hg21.txt | 2 --
2 files changed, 1 insertion(+), 3 deletions(-)
```

In the example below, we can see the difference along the master branch. The branch point between **B1** and **master** is still the same, node B. But the change is the node C we made on the master branch.



```
GitWorkshop/samples2$ git diff --compact-summary B1...master
hg17.txt | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)
```

This shows the difference between the branch-point of **master** and **B1** (node B) and the latest commit on **B1**.

Diff Symmetric Difference – Two Dots

Set symmetric difference is not the same as set subtraction. The *symmetric difference* between sets A and B is their union take away their intersection:

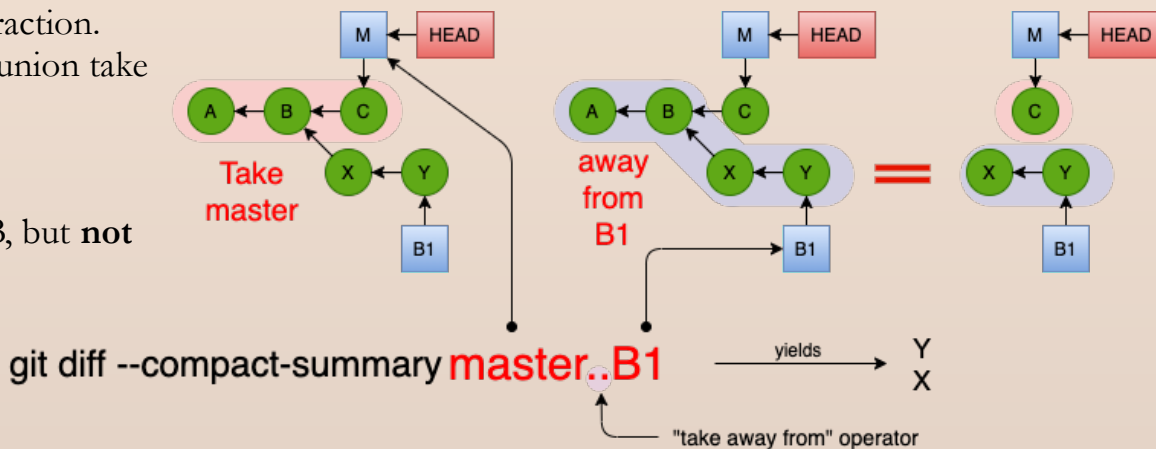
$$A \Delta B = A \cup B - A \cap B$$

The set of all elements that are in either in A or in B, but **not** in both.

Unlike displaying log messages, code changes have a "sign" or "direction". If you made a change by adding two lines to a file, then the reverse direction is removing those two lines from the file.

Despite the set theory term *symmetric difference*, the display of the change is actually **antisymmetric**. It has a distinct direction which is reversed when the order of the parameters to the operator are reversed.

In the diagram, the difference with double dots shows us how to go **from master to B1**. That means **undoing** commit C (changing "forty-two" back to "thirty-seven") and then **applying** commits X and Y.



```
GitWorkshop/samples2$ git diff --compact-summary master..B1
hg17.txt | 4 ++--
hg21.txt | 2 --
2 files changed, 2 insertions(+), 4 deletions(-)
GitWorkshop/samples2$ git diff --compact-summary B1..master
hg17.txt | 4 ++--
hg21.txt | 2 ++
2 files changed, 4 insertions(+), 2 deletions(-)
```

The opposite order means: start at Y, **undo** Y, **undo** X, and then **apply** C. Without the `--compact-summary` option, all the details of the changes are listed.

Exercise 7 - Merge

It's time to merge the **B1** branch to **master**. It's hard to remember what's happened on either of these two branches since the split occurred; and that's fairly realistic. So we'll start using the **git log** command technique to check what has occurred on each branch before starting the merge.

1. Summarize commits on **master** since **B1** split from **master**.

```
GitWorkshop/samples2$ git log --oneline B1..master
4878aee (HEAD -> master) Forty-two miles.
```

Just the one change on the one line.

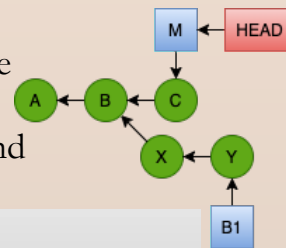
2. Summarize commits on **B1** since **B1** split from **master**.

```
GitWorkshop/samples2$ git log --oneline master..B1
81d60de (B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
```

3. With our memory refreshed, we proceed with the merge task. The **merge** command takes a single branch name for the source of the merge. **The target is always the current branch.** Since we want to merge into **master**, we must make **master** the current branch.

```
GitWorkshop/samples2$ git checkout master
Already on 'master'
Your branch is ahead of 'origin/master' by 1 commit.
```

As it happened, we were already on **master**. But it didn't hurt to check. We've confirmed our status depicted in the diagram on the right. The merge command itself is easy.



```
GitWorkshop/samples2$ git merge B1
Auto-merging hgl7.txt
Merge made by the 'recursive' strategy.
 hgl7.txt | 2 +-
 hgl21.txt | 2 --
 2 files changed, 1 insertion(+), 3 deletions(-)
```

4. Run the above git merge command. Git displays an editor window with the contents prepopulated as shown below.

```
Merge branch 'B1'
# Please enter a commit message to explain why this merge is necessary,
# especially if it merges an updated upstream into a topic branch.
#
# Lines starting with '#' will be ignored, and an empty message aborts
# the commit.
```

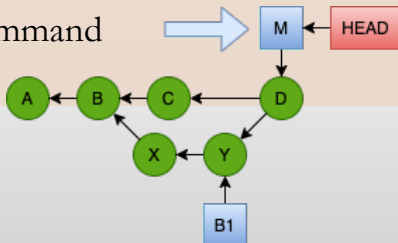
5. Replace the comment or accept it, save and quit. This creates the merge commit.

Merge Result

The merge commit is node **D** in the diagram below. Notice that commit **D** has two children: **C** and **Y**.

However, this is not apparent with the git log command we've been using.

```
GitWorkshop/samples2$ git log --online
6513089 (HEAD -> master) Merge branch 'B1'
4878aee Forty-two miles.
81d60de (B1) Songs to compositions.
c2e8b4e Pruned dead URL from Ch 21.
10f629d (origin/master, origin/HEAD) Added Python and fixed typos.
b83eb9b Initial version.
```



However, by adding the `--graph` option to the command, we can see the relationship of both children to the last commit.

```
GitWorkshop/samples2$ git log --graph --online
* 6513089 (HEAD -> master) Merge branch 'B1'
| \
| * 81d60de (B1) Songs to compositions.
| * c2e8b4e Pruned dead URL from Ch 21.
* | 4878aee Forty-two miles.
|/
* 10f629d (origin/master, origin/HEAD) Added Python and fixed typos.
* b83eb9b Initial version.
```

Without the graph option, the branches are flattened out. Of course, GUI tools make this visualization even nicer.

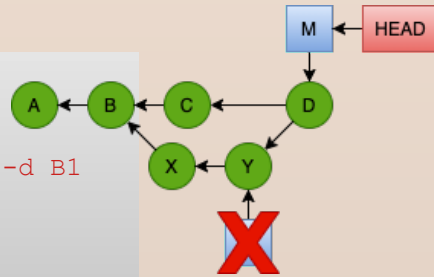
HEAD	Commit Hash	Commit Message	Date
master HEAD	6513089	Merge branch 'B1'	7/17/20, 1:52 PM
	4878aee	Forty-two miles.	6/19/20, 8:04 AM
B1	81d60de	Songs to compositions.	6/10/20, 3:07 PM
	c2e8b4e	Pruned dead URL from Ch 21.	6/10/20, 12:30 PM
origin/HEAD origin/master	10f629d	Added Python and fixed typos.	3/30/20, 9:02 PM
	b83eb9b	Initial version.	3/29/20, 1:15 PM

Displaying branch histories is one of the better advantages to using a GUI Git tool. Most IDEs have this capability built in or available as a free plugin.

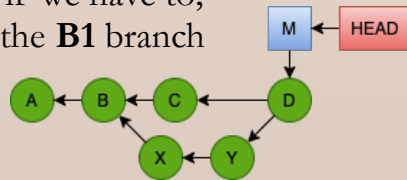
Deleting a Branch (pointer)

We're done with the **B1** branch; but the pointer is still hanging around. If we want to continue to work on a **B1** branch past this point, we probably want* to pick up contributions from the merge to **master**. To this end, we wish to remove the B1 pointer.

```
GitWorkshop/samples2$ git branch
  B1
* master
GitWorkshop/samples2$ git branch -d B1
Deleted branch B1 (was 81d60de).
GitWorkshop/samples2$ git branch
* master
```

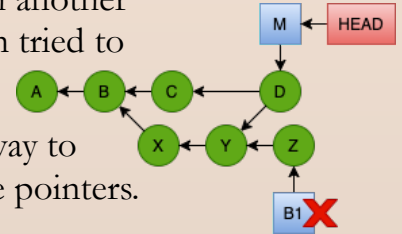


Git makes the deletion of the **B1** branch pointer easy. It was pointing to node **Y** and there was already another element pointing to **Y** (namely node **D**). So we can always reach node **Y** through **D** if we have to; we're not losing access to it by deleting the **B1** branch pointer.



What if we had continued with another commit on branch **B1** and then tried to delete **B1**?

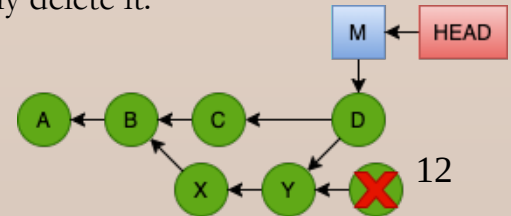
Then there would be no easy way to reach node **Z** through available pointers.



Git will recognize this and refuse the deletion with a warning that the branch is "not fully merged". Sometimes you still want to delete such a branch (for example, you want to discard any changes you made on the branch without a merge). Then you use the same command with a capital **D**.

```
git branch -D B1
```

This will delete the branch pointer regardless of whether it had been merged. After this, node **Z** is essentially unreachable since there is no path to it through available pointers. Git will eventually delete it.



* There are times when we want to continue working on B1 without the master contributions. But doing this for too long increases the risk of complicating subsequent merges.

Merge Conflicts

We are going to introduce some changes in our files on different branches that **conflict**. In this context, it means that different branches change the same line of the same file in different ways. We saw earlier that neither

- changing different files, nor
- changing the same file on different lines

incur a conflict. The **git merge** command will merge these deterministically.

In the following exercise we edit two files: a plain text file and a small Python program.

```
1 # This file contains mappings.
2 #
3 a1 - 20
4 a2 - 43
5
6 b1 - 39
7 b2 - 34
8 b3 - 44
9
10 c1 - 45
11 c2 - 19
```

file1.txt

To keep things distinct from Part 1, we'll create a new directory, `sample3`, with these files and create a new repository.

```
GitWorkshop/samples3$ ls
file1.txt  file2.py
GitWorkshop/samples3$ git init
Initialized empty Git repository in GitWorkshop/samples3/.git/
GitWorkshop/samples3$ git add .
GitWorkshop/samples3$ git commit -m "Initial version."
[master (root-commit) 4fff5a8] Initial version.
 2 files changed, 22 insertions(+)
 create mode 100644 file1.txt
 create mode 100644 file2.py
GitWorkshop/samples3$ git log --oneline
4fff5a8 (HEAD -> master) Initial version.
```

```
1 def print_usage():
2     usage = """Usage: addAudit.py [-f] [-v] <filename ...>
3         -f - overwrite when duplicate key encountered
4         -v - verbose
5         <filename ...> the name of at least one audit file."""
6
7     print(usage)
8
9 print_usage()
```

file2.py

Exercise 8 – Branch B2

Create a branch **B2** and edit the two files.

1. Create a new branch **B2**.

```
GitWorkshop/samples3$ git checkout -b B2
Switched to a new branch 'B2'
```

2. Edit `file1.txt` as shown on the right.

- a. **Line 6:** Add ",41"
- b. **Line 7:** Change "34" to "36"
- c. **Line 8:** Add a blank space after 44.
- d. **Line 10:** Change "45" to "55"
- e. **Line 11:** Change "19" to "29"

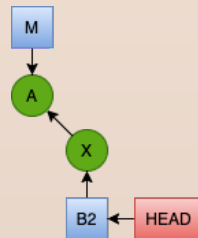
3. Edit `file2.py` by adding a blank space to each line of the `print_usage` function.

4. Check your work with `git diff`.

5. Add the changes to the staging area.

6. Commit with message "B2 changes."

This creates a new commit **X** referenced by branch **B2** as shown in the figure above.



```
1 # This file contains mappings.
2 #
3 a1 - 20
4 a2 - 43
5
6 b1 - 39,41
7 b2 - 346
8 b3 - 44
9
10 c1 - 455
11 c2 - 1929
```

file1.txt

```
1
2 def print_usage():
3     usage = """Usage: addAudit.py [-f] [-v] <filename ...>
4     -f - overwrite when duplicate key encountered
5     -v - verbose
6     <filename ..> the name of at least one audit file."""
7
8     print(usage)
9
10 print_usage()
```

file2.py

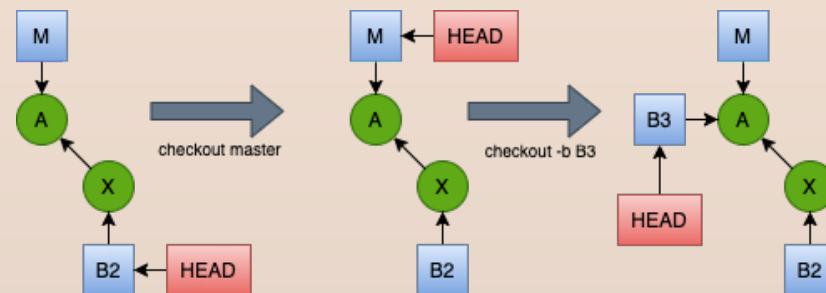
Note that some changes are not as easy to see with the **git diff** command.

Extra Credit: The edit to `file2.py` may seem peculiar.
What is a common cause for this?

Exercise 8 – Branch B3

7. Now we're going switch to a new branch **B3** starting from commit **A** just like **B2** did.

```
GitWorkshop/samples3$ git checkout master
Switched to branch 'master'
GitWorkshop/samples3$ git checkout -b B3
Switched to a new branch 'B3'
```



8. Edit file1.txt.

- a. **Line 3:** change "20" to "30".
- b. **Line 4:** change "43" to "53".
- c. **Line 6:** add ",40".
- d. **Line 7:** change "34" to "35".



```
1 # This file contains mappings.
2 #
3 a1 - 230
4 a2 - 453
5
6 b1 - 39,40
7 b2 - 345
8 b3 - 44
9
10 c1 - 45
11 c2 - 19
```

file1.txt

9. Edit file2.py. Change the triple-quoted string to a series of print statements.



```
1 def print_usage():
2     print("Usage: addAudit.py [-f] [-v] <filename ...>")
3     print("    -f - overwrite when duplicate key encountered")
4     print("    -v - verbose")
5     print("    <filename ..> the name of at least one audit file.")
6
7 print_usage()
```

file2.py

10. Check your work with git diff.
11. Add the changes to the staging area.
12. Commit with message "B3 changes."

Exercise 8 – Branch master

After the B3 commit, we have the branch configuration shown to the right. Since the target of a merge is always the current branch, we're going to change back to the **master** branch. Then merge each of **B2** and **B3**.

13. Change to the **master** branch.

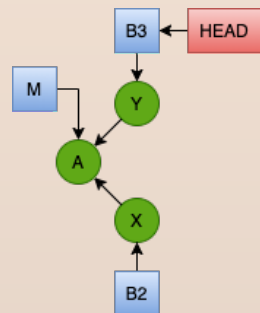
```
git checkout master
```

This moves HEAD to the master branch.

14. Merge **B2**

```
git merge B2
```

This should be a fast-forward merge which simply advances the master branch to **B2**.



15. Finally, issue the command to merge **B3**. This is where the fireworks start.

```
GitWorkshop/samples3$ git merge B3
Auto-merging file2.py
CONFLICT (content): Merge conflict in file2.py
Auto-merging file1.txt
CONFLICT (content): Merge conflict in file1.txt
Automatic merge failed; fix conflicts and then commit the result.
```

What just happened here?

Git auto-merges on a line-by-line basis. When the same line is changed in two different ways, Git places "merge markers" around those lines to indicate two incompatible changes were made. They must be manually resolved.

The changes are easy to spot. They are bounded by the markers

```
<<<< HEAD – beginning of master
```

```
===== – separates versions
```

```
>>>>> B2 – end of B3 version
```

Resolving the merge conflict amounts to choosing one or the other of these versions to keep.

```
1 # This file contains mappings.
2 #
3 a1 - 30
4 a2 - 53
5
6 <<<<<<< HEAD
7 b1 - 39, 41
8 b2 - 36
9 b3 - 44
10 =====
11 b1 - 39, 40
12 b2 - 35
13 b3 - 44
14 >>>>>>> B3
15
16 c1 - 55
17 c2 - 29
```

This is the HEAD (i.e. master) version of the lines (originally B2)

This is the B3 (i.e. incoming) version of the lines

Exercise 8 – Resolve Conflict 1

We have to decide, based on a larger perspective, how to resolve the conflicts. We could

- pick the left side (**master** branch version)
- pick the right side (**B3** branch version)
- choose something completely different from either side based on some knowledge we might have.

These decisions are carried out in the following manner.

- Edit the lines** within the merge markers based on your decisions
- Delete the merge marker** lines and save the file.
- Add the file** to the Git staging area.

This last step is how Git knows when we've completed the merge activity for the file. We repeat steps a, b and c for each file in which a merge conflict occurred.

Looking at `file1.txt` in the previous slide, things aren't that bad. The top (only changed on **B2**) and the bottom (only changed on **B3**) were auto-merged. Only the middle third, where both **B2** and **B3** changed lines, requires resolution.

16. Edit the lines of `file1.txt`.

- For entry `b1`, branch **B2** added `41` while branch **B3** added `40` to the value. Let's make the decision to add both so that line 7 has `b1 - 39, 40, 41`. Notice how we're implicitly making lines 7 – 9 our "definitive copy."
- For entry `b2`, branch **B2** changed the value to `36` which branch **B3** changed it to `35`. Let's decide to keep `36` so that line 8 remains unchanged.
- Entry `b3` is tricky. It looks the same in both lines 9 and 13. Recall that branch **B2** erroneously added a space at the end of the line. In this case, we wish to accept the **B3** line 13 which left the line unchanged. Remove the last space on line 9.

```
6 <<<<<<<•HEAD␣  
7 b1•-•39, 41␣  
8 b2•-•36␣  
9 b3•-•44•␣  
10 =====␣  
11 b1•-•39, 40␣  
12 b2•-•35␣  
13 b3•-•44␣  
14 >>>>>>>•B3␣
```



```
6 <<<<<<<•HEAD␣  
7 b1•-•39, 40, 41␣  
8 b2•-•36␣  
9 b3•-•44␣  
10 =====␣  
11 b1•-•39, 40␣  
12 b2•-•35␣  
13 b3•-•44␣  
14 >>>>>>>•B3␣
```

Exercise 8 – Resolve Conflicts 2

17. Delete the merge markers. Lines 7 – 9 are now in the form we want to keep. We can delete lines 11 – 13 as well as the merge markers on lines 6, 10 and 14. Then save the file.

```
6 <<<<<< HEAD
7 b1 - 39,40,41
8 b2 - 36
9 b3 - 44
10 =====
11 b1 - 39,40
12 b2 - 35
13 b3 - 44
14 >>>>>> B3
```



```
6 b1 - 39,40,41
7 b2 - 36
8 b3 - 44
```

```
1 def print_usage():
2 <<<<<< HEAD
3     usage = """Usage: addAudit.py [-f] [-v] <filename ...>
4         -f - overwrite when duplicate key encountered
5         -v - verbose
6         <filename ..> the name of at least one audit file."""
7
8     print(usage)
9 =====
10    print("Usage: addAudit.py [-f] [-v] <filename ...>")
11    print("    -f - overwrite when duplicate key encountered")
12    print("    -v - verbose")
13    print("    <filename ..> the name of at least one audit file.")
14 >>>>>> B3
15
16 print_usage()
```

18. Add file1.txt to the Git staging area.

```
git add file1.txt
```



19. Perform the same merge steps a, b and c for file2.py. This is a simpler case where we wish to only accept the **B3** version. The **B2** version had simply added an extra space on each line, which is not uncommon for some editors. Simply delete lines 2 – 9 and line 14. Then save the file and add it to the staging area.

```
1 def print_usage():
2     print("Usage: addAudit.py [-f] [-v] <filename ...>")
3     print("    -f - overwrite when duplicate key encountered")
4     print("    -v - verbose")
5     print("    <filename ..> the name of at least one audit file.")
6
7 print_usage()
```

Exercise 8 - Commit

20. With both file conflicts resolved and added to the Git staging area, we can now create the merge commit. It should be like any other commit; the hard work is over.

```
git commit -m "Merge branch B3"
```

21. Verify the branch activity using **git log**.

```
GitWorkshop/samples3$ git log --oneline --graph
*   55a29e6 (HEAD -> master) Merge branch B3
| \
| * 17c3dc5 (B3) B3 changes.
* | fd4d195 (B2) B2 changes.
|/
* cce1b27 Initial version.
```

And that's all there is to it. It's just a matter of recognizing the merge markers and resolving the merge the proper way.

Note the merge-marker technique only works with text files. This technique does not work for binary files, such as images, executables, compressed archives, etc. Teams must take care not to change these files concurrently. They cannot easily be merged.

Remote Repositories – Terms

A Git *clone* is another copy of the current repository.

- Sometimes it's called a *remote repository*. It's remote in the sense that it's in a separate directory. It doesn't have to be on a separate machine. We saw examples early in Part 1 where we created multiple remote repositories on the same machine with the `clone` command. The `clone` command is the typical way to create a remote repository. The term *clone* can refer to the remote repository, or it can refer to the Git command.
- A clone is not a strict copy. By default, the `clone` command only copies commits reachable by the master branch. Additional branches may be cloned through additional commands.
- In this presentation, we'll use the term **clone** as
 - a Git command `clone` using monospace font,
 - a noun – referring to a clone of a repository
 - a verb – the act of creating a clone with `clone`.

A Git *remote* is an alias for a reference to a remote repository location.

- It saves from having to type the full remote URL each time it is referenced.
- It simply maps a name to a URL. It does not confirm or deny the existence of a repository at that URL.

A repository created as a clone is automatically configured with a remote named **origin** referencing the source repository.

```
GitWorkshop$ git clone https://github.com/lacounty-isab/workshops
myrepo
Cloning into 'myrepo'...
Resolving deltas: 100% (280/280), done.
GitWorkshop$ cd myrepo
GitWorkshop/myrepo$ git remote -v
origin https://github.com/lacounty-isab/workshops (fetch)
origin https://github.com/lacounty-isab/workshops (push)
```

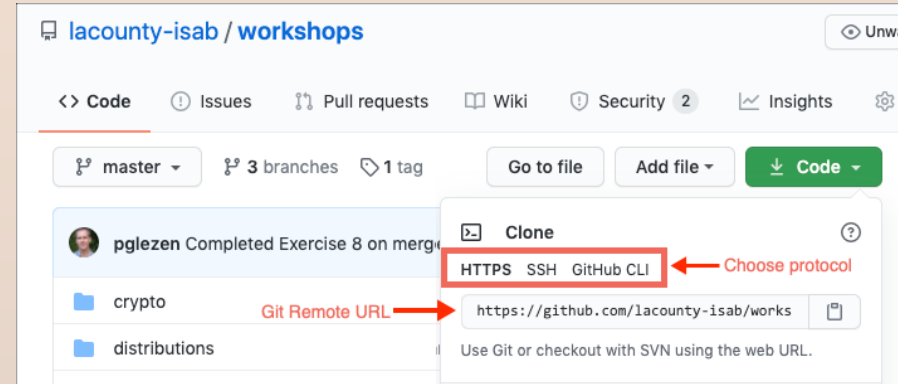
The remote name, `origin`, is configured for *fetch* and *push* (more on these commands later). This is done for convenience since one often wishes to perform synchronization operations after a clone.

No remote entries initially exist for new repositories.

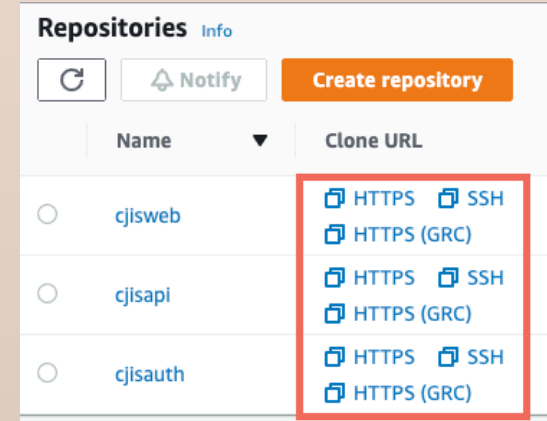
Remote Command

The `git remote` command is used to **add**, **view**, **rename** and **remove** remote entries. By default the command will list the names of remotes. Add `-v` to see the full entry. Below are examples using the repository we clone in the previous slide.

```
GitWorkshop/myrepo$ git remote
origin
GitWorkshop/myrepo$ git remote -v
origin      https://github.com/lacounty-isab/workshops (fetch)
origin      https://github.com/lacounty-isab/workshops (push)
GitWorkshop/myrepo$ git remote rename origin github
GitWorkshop/myrepo$ git remote -v
github      https://github.com/lacounty-isab/workshops (fetch)
github      https://github.com/lacounty-isab/workshops (push)
GitWorkshop/myrepo$ git remote remove github
GitWorkshop/myrepo$ git remote -v
GitWorkshop/myrepo$ git remote add origin
https://github.com/lacounty-isab/workshops
GitWorkshop/myrepo$ git remote -v
origin      https://github.com/lacounty-isab/workshops (fetch)
origin      https://github.com/lacounty-isab/workshops (push)
```



Most Git hosting services provide the remote URL on the repository page. The figure above shows **GitHub**. The figure to the right shows **AWS CodeCommit**.



Remotes

Protocols

Git supports several protocols for communicating between a Git server listener and a Git client.

- **local** – In this case there is no network activity; the Git client works with both repositories on your workstation.
- **git** – This is a proprietary protocol that is rarely used anymore. The advantage is that it works with little configuration. The drawbacks are that it is neither secure nor firewall friendly.
- **ssh** – This tunnels the Git protocol over SSH. This has the benefit of encrypting the connection and it is very firewall friendly. The drawback is that it requires each client to install an SSH certificate on the server.
- **https** – This mode acts like an HTTPS server. The user must be authenticated on the first request. But usually the credentials can be cached and reused later.

These days, Git hosting services take care of the security configurations, making the bottom two options the most common.

Bare Repositories

- We briefly discussed bare repositories in Part 1. A bare repository has no working copy, just the repository itself. All server copies are bare repositories since you can only push a commit to a bare repository. You can pull from a non-bare; but it's an error to push to one.
- A naming convention for directories containing a bare repository is to add `.git` to the end. It's unusual to add extensions to folder names, so this makes the folder's use clear.
- Even though most Git hosting services implement this convention, they still allow us to leave the `.git` off the end of our remote URL definitions. However, for the local folders we use in our exercises, the Git client will be more strict.

Remote Branch References

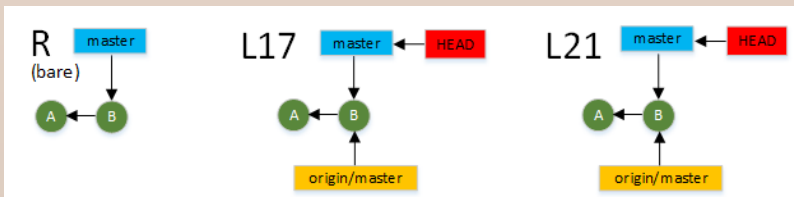
We reference a local branch simply by its name. Remote branches are prefaced by their remote name. If the remote name is `origin`, then a reference to the master branch would be `origin/master`.

Tracking Branches

A **tracking branch** tracks a branch in a remote repository. It's important to understand the limitations to remote branch tracking. This will be one of the goals of the next exercise.

Exercise 9 – Preparation

We're going to simulate a remote repository for the next exercise. But it should be emphasized that the "simulation" is only in regard to the network connection. In every other way, it will be a legitimate remote. Git does not treat remote repositories on separate machines differently than it does remote repositories on the local machine.



1. Clone our sample repository to a new directory samples8.

```
GitWorkshop$ git clone  
https://github.com/lacounty-isab/gitwkspe6 samples8
```

This creates the source for our scenario.

2. Create a bare repository to represent the shared server repository.

```
GitWorkshop$ git clone --bare samples8 sR
```

The "samples Remote" will hence be known as **sR**.

3. Clone sR to sL17 and sL21.

```
GitWorkshop$ git clone sR sL17  
Cloning into 'sL17'...  
done.  
GitWorkshop$ git clone sR sL21  
Cloning into 'sL21'...  
done.
```

After these three preparation steps, you should have three new directories: sR, sL17 and sL21, corresponding to the three DAGs shown on the left.

Exercise 9 – L17

First we simulate the user at repository sL17.

4. Change to the sL17 directory.
5. Change "songs" to "compositions" on line 17 of hg17.txt like we did before.
6. Add and commit the change.

```
GitWorkshop$ cd sL17
GitWorkshop/sL17$ vi hg17.txt
GitWorkshop/sL17$ git add .
GitWorkshop/sL17$ git commit -m "Songs to Compositions in hg17.txt"
GitWorkshop/sL17$ git log --oneline
8e5d42c (HEAD -> master) Songs to Compositions in hg17.txt
d0d9eea (origin/master, origin/HEAD) Added Python and fixed typos.
79647d7 Initial version.
```

The listing above shows three commits in the log in reverse chronological order. These represent:

- **C** – 8e5d42c – the commit we just made
- **B** – d0d9eea – the newest commit on the server
- **A** – 79647d7 – the initial commit

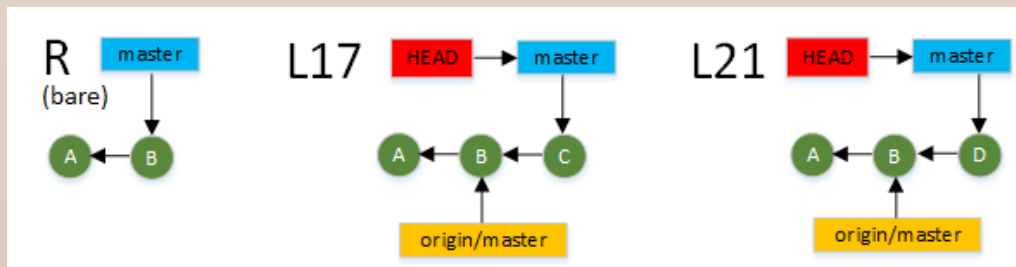
List the branch pointer locations.

```
GitWorkshop/sL17$ git branch -av
* master          8e5d42c [ahead 1] Songs to
Compositions in hg17.txt
remotes/origin/HEAD -> origin/master
remotes/origin/master d0d9eea Added Python and fixed typos.
```

Verify that

- master – points to **C**
- origin/master – the tracking branch still points to **B**

Note: We're ignoring the right-side of the figure for now.



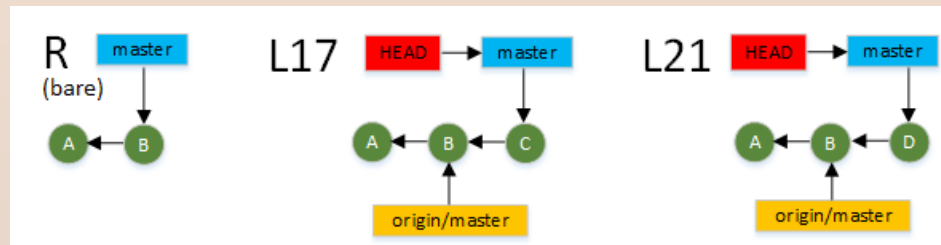
Exercise 9 – L21

Now let's assume **L21** also creates a commit in their local repository. Then we get the figure on the right. We're not going to simulate **L21** yet; instead, we're going to stick with **L17** and show why things go wrong.

The next step for **L17** is to push the commit to the remote **R**. It's important to understand that we cannot push to a remote unless the **remote tracking branch** is consistent with the **remote branch**. In the diagram to the right,

- **remote tracking branch** is the yellow box labelled `origin/master` in **L17**.
- **remote branch** is the blue box `master` in **R**.

Presently, they both point to commit **B**, which makes them consistent. But this is not obvious to **L17**. Perhaps someone pushed a commit to **R** since we last checked. In order to check, **L17** should refresh the remote tracking branch to make sure it is accurate. Then push the update.



7. Refresh tracking branch.

```
GitWorkshop/sL17$ git fetch origin master
From /Users/pglezen/lac/sessions/GitWorkshop/sR
 * branch          master      -> FETCH_HEAD
GitWorkshop/sL17$ git log --oneline
8e5d42c (HEAD -> master) Songs to Compositions in hgl7.txt
d0d9eea (origin/master, origin/HEAD) Added Python and fixed typos.
79647d7 Initial version.
```

8. Once we see the branch is consistent, we push the change.

```
GitWorkshop/sL17$ git push origin master
... d0d9eea..8e5d42c master -> master
GitWorkshop/sL17$ git log --oneline
8e5d42c (HEAD -> master, origin/master, origin/HEAD) Songs to
Compositions in hgl7.txt
d0d9eea Added Python and fixed typos.
79647d7 Initial version.
```

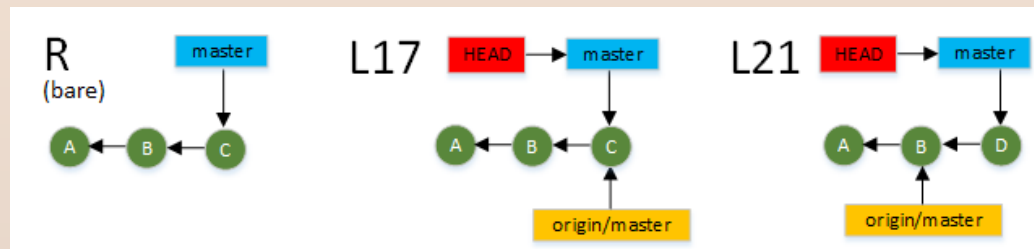
Exercise 9 – First Commit

But this doesn't bid well for **L21**. An attempt to refresh from **R** will show more than just that **L21** is out of date. It will show that the master sequence of **R** (**C**, **B**, **A**) is incompatible with the master sequence of **L21** (**D**, **B**, **A**). At this point **L21** is stuck and can no longer work with the master branch.

This is what happens when two developer try to commit to the same branch without proper synchronization/communication. In fact, you don't even need two developers. The same developer with a clone of the repository on different workstations can encounter the same problem.

This problem can be reduced by remembering to update the remote tracking branch with a fetch before each commit.

But there is still a possiblity that in the small amount of time between the tracking branch update and the commit, another commit will be made. The true way to avoid this problem is to limit master branch commits to a single person or to a small group of people who communicate well. Everyone else commits using different branches that are merged to master. This eliminates the race conditions.



The rest of this exercise will be demonstration of this practice. We'll go back in time and designate **L17** as the owner of the master branch. As such, **L17** is also responsible for merging other branches into master.

9. Change to the **L21** clone directory and create a new **i42** branch.

```
GitWorkshop/sL17$ cd ../sL21
GitWorkshop/sL21$ git branch -a
* master
  remotes/origin/HEAD -> origin/master
  remotes/origin/master
GitWorkshop/sL21$ git checkout -b i42
Switched to a new branch 'i42'
GitWorkshop/sL21$ git branch -a
* i42
  master
  remotes/origin/HEAD -> origin/master
  remotes/origin/master
```

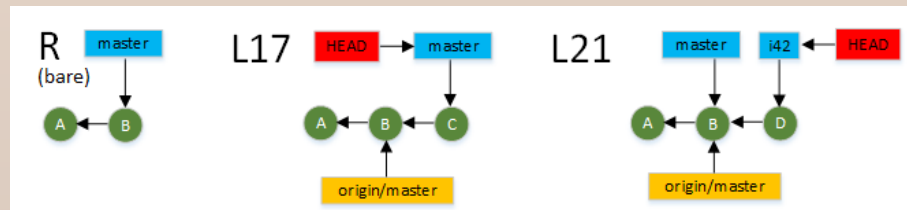
Exercise 9 – Commit to L21

10. Remove the first two lines from `hg21.txt` (which should be a URL and blank line).

11. Add the change to the staging area and commit.

```
GitWorkshop/sL21$ git add hg21.txt
GitWorkshop/sL21$ git commit -m "Removed two lines."
[i42 02a6929] Removed two lines.
1 file changed, 2 deletions(-)
GitWorkshop/sL21$ git log --oneline
02a6929 (HEAD -> i42) Removed two lines.
d0d9eea (origin/master, origin/HEAD, master) Added Python
and fixed typos.
79647d7 Initial version.
```

We're now at the point where both developers have something to commit on top of the **B** commit. **L17** committed to local master as before. But **L21** committed to local `i42`.

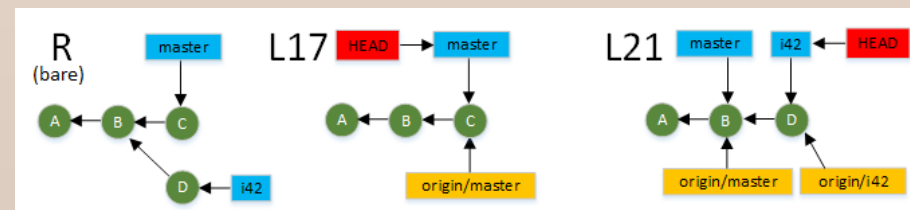


Now each developer pushes his local branch to the remote server. The L17 push was done in Step 8.

12. We'll push the `i42` branch below.

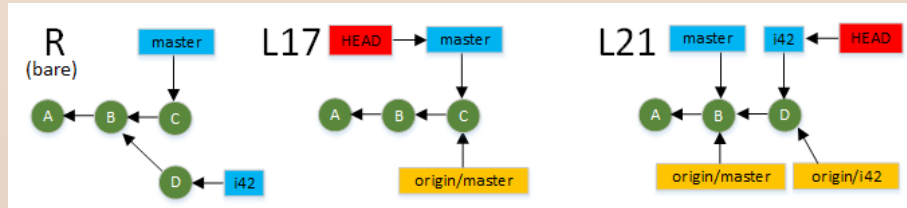
```
GitWorkshop/sL21$ git push origin i42
...
To GitWorkshop/sR
 * [new branch]      i42 -> i42
GitWorkshop/sL21$ git branch -a
* i42
  master
remotes/origin/HEAD -> origin/master
remotes/origin/i42
remotes/origin/master
```

Notice that the remote tracking branch `origin/i42` was created by the first `git push` for the branch. It tracks the remote `i42` branch in **R**.



Exercise 9 – Fetch

Another important point is that the local master branch point has not changed in the **L21** repository. It's outdated (which is fine); but not in conflict.



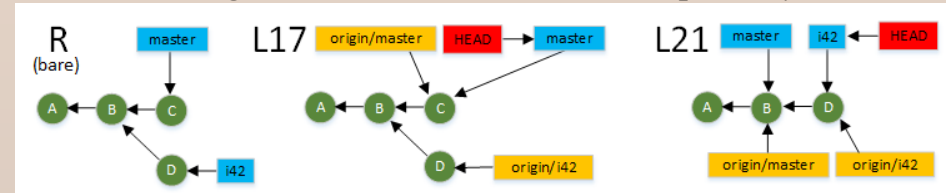
Note also that at this point in time all three repositories, the two local ones and the shared remote, have different sets of commits. This is not a problem so long as branches are used in this way.

At this point the developer for **L21** informs the master owner (**L17**) that branch `i42` is available in the shared remote for merging into master.

13.L17 fetches the `i42` branch.

```
GitWorkshop/sL21$ cd ../sL17
GitWorkshop/sL17$ git branch -a
* master
  remotes/origin/HEAD -> origin/master
  remotes/origin/master
GitWorkshop/sL17$ git fetch origin i42
...
From GitWorkshop/sR
 * branch          i42          -> FETCH_HEAD
 * [new branch]    i42          -> origin/i42
GitWorkshop/sL17$ git branch -a
* master
  remotes/origin/HEAD -> origin/master
  remotes/origin/i42
  remotes/origin/master
```

The listing above and diagram below shows how the additional remote tracking branch is added to the L17 repository.



It's important to understand that while the `i42` branch code has been added to the repository, is nowhere to be seen in our working copy, which still reflects master.

Exercise 9 – Another Commit

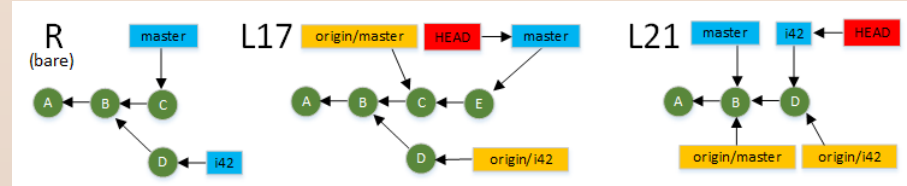
The **L17** developer has fetched the `i42` branch, but will not work with it yet; and that's perfectly fine. Fetching the content of other branches does **not** disrupt our working copy activity.

14. Still in the **L17** repository, edit `hg17.txt`: change the text on Line 10 from `thirty-seven` to `forty-two`.

```
-bunkers some thirty-seven miles from the stage,  
+bunkers some forty-two miles from the stage,
```

15. Add and commit the change to `master`.

```
GitWorkshop/sL17$ git commit -m "37 to 42"  
[master 01a19f8] 37 to 42  
1 file changed, 1 insertion(+), 1 deletion(-)  
GitWorkshop/sL17$ git log --oneline  
01a19f8 (HEAD -> master) 37 to 42  
8e5d42c (origin/master, origin/HEAD) Songs to  
Compositions in hg17.txt  
d0d9eea Added Python and fixed typos.  
79647d7 Initial version.
```



The last commit is **E** on the diagram. The **L17** local `master` is now one commit ahead of the remote tracking branch `master`. The log listing in Step 15 and the figure both show the four commits of the **L17** local `master` branch.

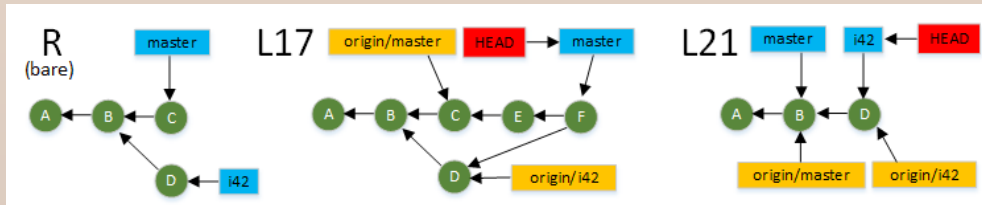
Exercise 9 – Merge

Developer **L17** is now ready to merge the changes from branch `i42`. Note that we have no local `i42` branch. All we have is the remote tracking branch. But that's fine since we're only reading from it as the source of the merge.

16. Merge the `i42` branch from `origin/i42` into `master`.

```
GitWorkshop/sL17$ git branch -a
* master
  remotes/origin/HEAD -> origin/master
  remotes/origin/i42
  remotes/origin/master
GitWorkshop/sL17$ git merge origin/i42 -m "Merged i42 to master."
Merge made by the 'recursive' strategy.
 hg21.txt | 2 --
 1 file changed, 2 deletions(-)
```

The comment is included in the merge command for this slide; it's common to leave it off and enter it interactively.



The DAG for **L17** is reflected in the `git log` command if we include the `--graph` option.

```
GitWorkshop/sL17$ git log --oneline --graph
* 82b8b8f (HEAD -> master) Merged i42 to master.
|\
| * 02a6929 (origin/i42) Removed two lines.
* | 01a19f8 37 to 42
* | 8e5d42c (origin/master, origin/HEAD) Songs to
  | Compositions in hg17.txt
  |/
* d0d9eea Added Python and fixed typos.
* 79647d7 Initial version.
```

Be sure you can identify the following in both the figure the left and the log output above:

- The merge commit **F** (82b8b8) is the new HEAD of the `master` branch.
- The remote tracking branch for `master` is behind local for **L21**, but still accurately reflects the state of the remote **R**.
- In fact, this local activity has resulted in **no changes** to any remote tracking branches.

Exercise 9 – Push

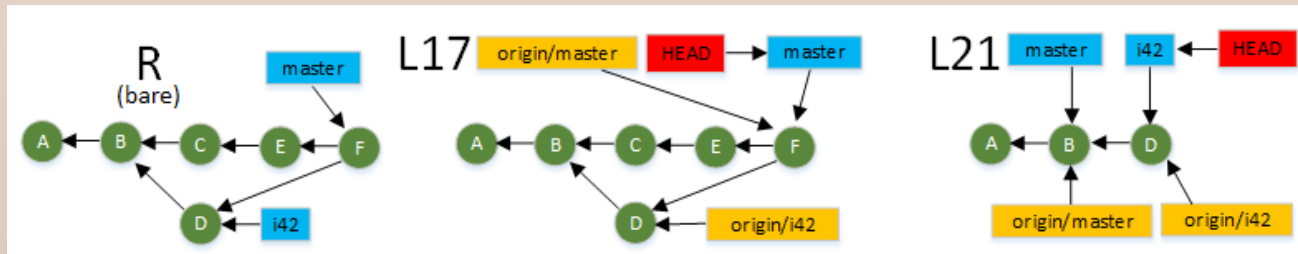
17. **L17** pushes the result of the merge to the shared remote **R**.

```
GitWorkshop/sL17$ git push origin master
...
GitWorkshop/sL17$ git log --oneline --graph
* 82b8b8f (HEAD -> master, origin/master, origin/HEAD) Merged i42 to master.
|\
| * 02a6929 (origin/i42) Removed two lines.
* | 01a19f8 37 to 42
* | 8e5d42c Songs to Compositions in hg17.txt
|/
* d0d9eea Added Python and fixed typos.
* 79647d7 Initial version.
```

The remote **R** now has the two new commits from **L17** along with the fact that commit **F** is a merge of commits **D** and **E**.

It is now the responsibility of **L17** to notify **L21** (and anyone else) of the availability of the merge.

From the **L17** log, the `git push` is reflected only in the movement of the tracking branch pointer.



Exercise 9 – Fetch from Remote

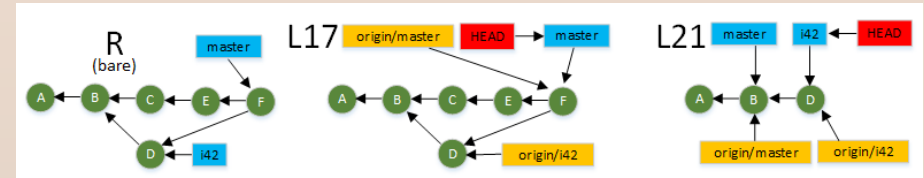
The **L21** developer's work has been merged into the master branch by the **L17** developer and pushed to the shared remote repository **R**. But this is still not visible to the **L21** developer until refreshed from the remote master.

18. Change to the **L21** repository and fetch the updated master branch from **R**.

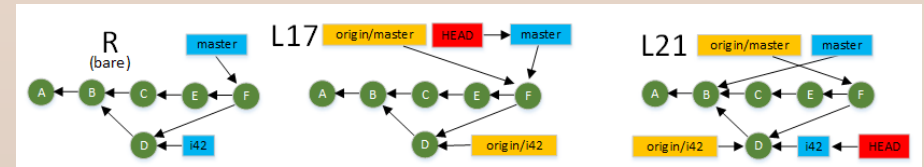
```
GitWorkshop/sL17$ cd ../sL21
GitWorkshop/sL21$ git log --oneline --graph
* 02a6929 (HEAD -> i42, origin/i42) Removed two lines.
* d0d9eea (origin/master, origin/HEAD, master) Added Python
and fixed typos.
* 79647d7 Initial version.
GitWorkshop/sL21$ git fetch origin master
. . . . .
From GitWorkshop/sR
 * branch      master      -> FETCH_HEAD
   d0d9eea..82b8b8f master  -> origin/master
GitWorkshop/sL21$ git log --oneline --graph
* 02a6929 (HEAD -> i42, origin/i42) Removed two lines.
* d0d9eea (master) Added Python and fixed typos.
* 79647d7 Initial version.
```

Note once again that the fetch did **not** disrupt any work **L21** might have been doing. The logs before and after the fetch show the same entries.

However, close inspection of the log outputs show that some branch pointers have moved around. Let's compare the figures. Below is the situation **before** the fetch, reproduced from the last slide.



And this is the situation **after** the fetch.



The **L21** situation looks a lot different in the figure. But it's consistent with the log output. The current branch is still **i42** and **origin/master** is no longer reachable from **i42**; though local **master** (which is now 3 commits behind) is still reachable. Both log output shows **D** → **B** → **A**.

Exercise 9 – Merge to Local Master

The previous slide showed **L21** could fetch all the merge activity without disrupting local activity. But now it's time to refresh the merge to our local master. As always with a merge, the first step is to change to the target branch.

19. Change to local master branch.

```
GitWorkshop/sL21$ git checkout master
Switched to branch 'master'
Your branch is behind 'origin/master' by 4 commits, and can be fast-forwarded.
(use "git pull" to update your local branch)
```

It tells us master is behind by 4 commits. But the good news is that it can be *fast-forwarded*; otherwise we have to "merge the merge", which we wouldn't do because we're not suppose to be contributing commits directly to master.

20. Merge from the remote tracking branch.

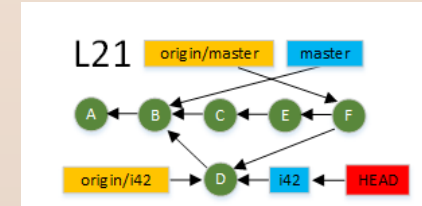
```
GitWorkshop/sL21$ git merge origin/master
Updating d0d9eea..82b8b8f
Fast-forward
```

All merges to master should be fast-forward for developers who do not commit directly to master.

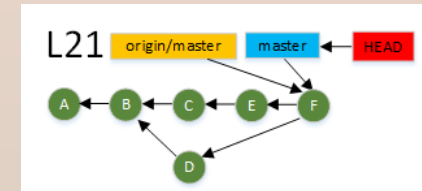
21. Remove the local and remote i42 branch pointer.

```
GitWorkshop/sL21$ git branch -d i42
Deleted branch i42 (was 02a6929).
GitWorkshop/sL21$ git branch -d -r origin/i42
Deleted remote-tracking branch origin/i42 (was 02a6929).
```

Note that deleting a remote branch requires the `-r` flag.



- Checkout master
- Merge origin/master
- Delete i42



Pull

If **L21** had already been on the `master` branch to begin with, the two steps

1. `git fetch origin master`
2. `git merge origin/master`

could have been combined into a single command.

```
git pull origin master
```

The `git pull` command combines the `fetch` and `merge` commands into a single operation.

End of Part 2

- Still to go (Part 3)
 - Tags
 - Hosting Services
 - GitHub
 - AWS CodeCommit
 - Stashing
 - Commit Signing
 - Review your GPG notes
 - Have your private key ready

Part 2 was the most advanced of the this three part workshop. Part 3 contains items that are useful in day-to-day usage of Git.