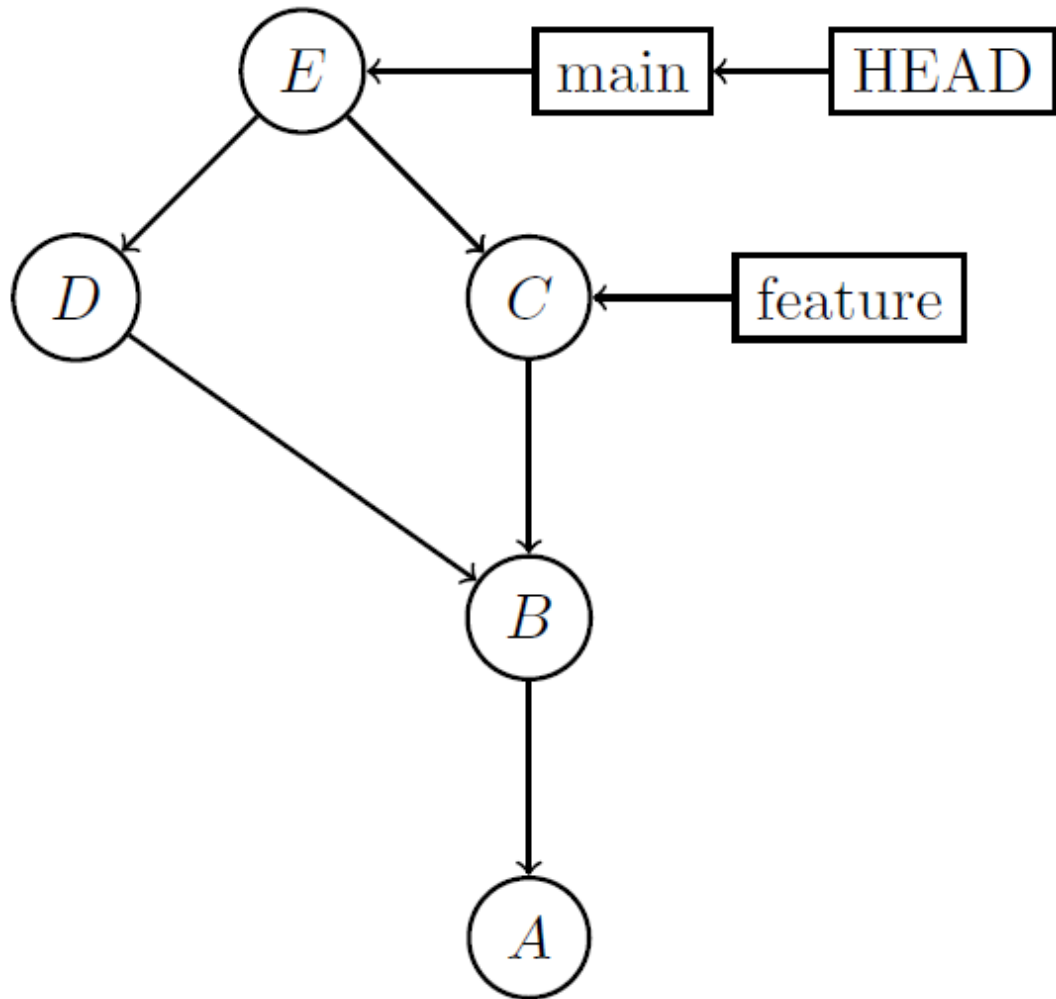


Understanding Git

DANIEL MOSCOE

DATA SCIENCE IN CONTEXT

3/3/2021



A consistent model for a Git repo: the digraph

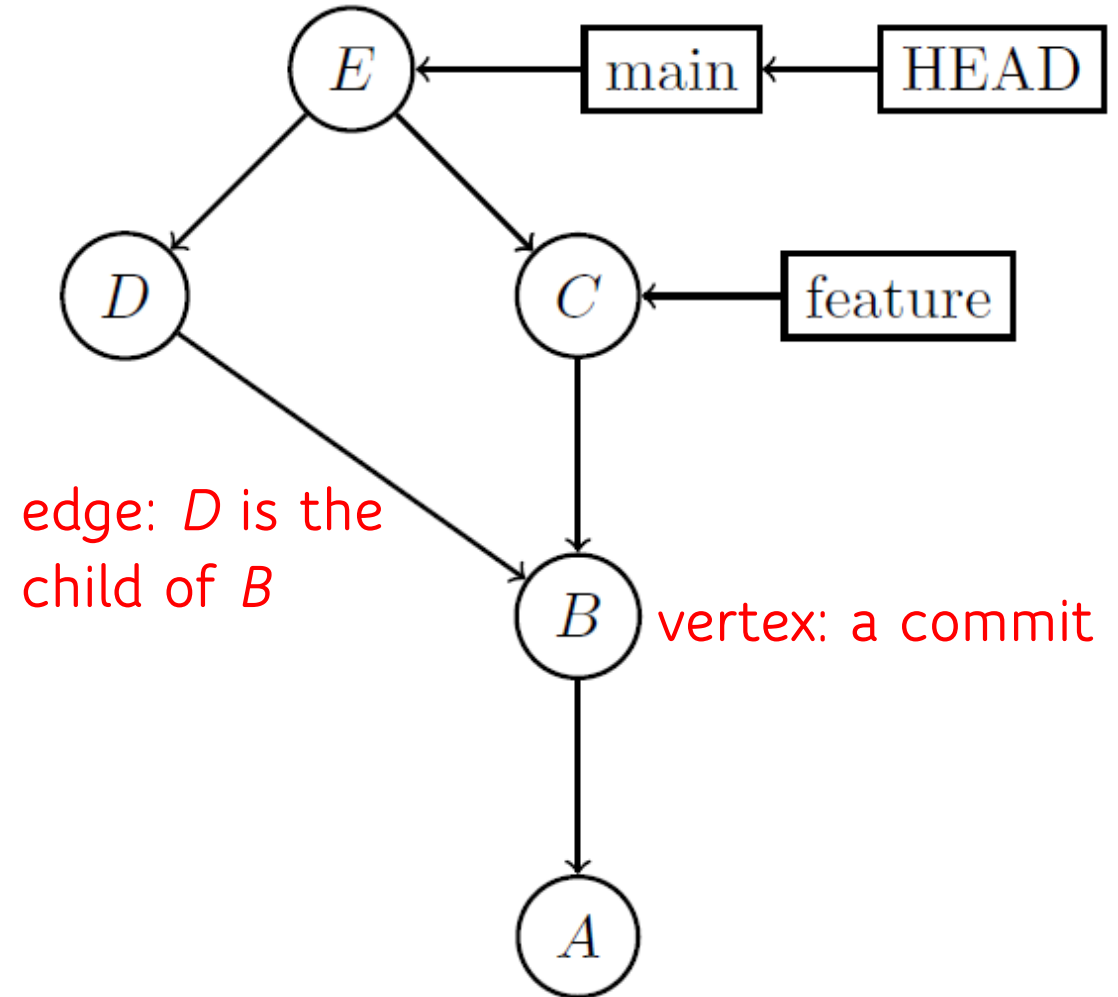
- ✓ A directed graph (digraph) is a collection of *vertices* and *edges*.
- ✓ Git / Hub are tools for (collaborating on) constructing and editing a directed graph.
- ✓ *To understand Git, compare multiple representations of the repo digraph.*

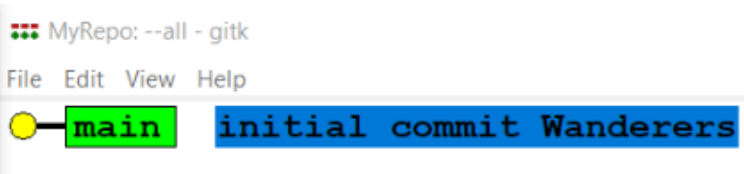
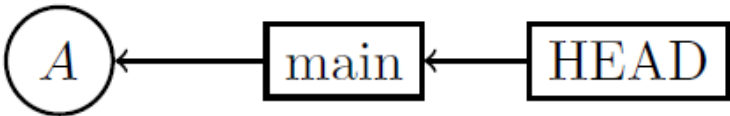
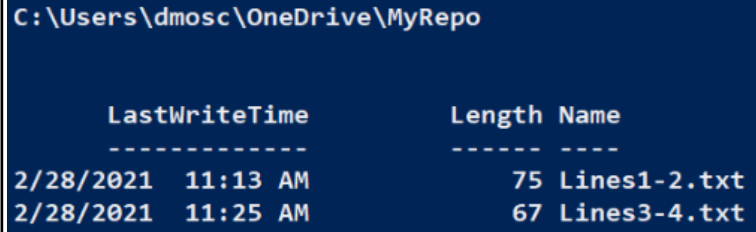
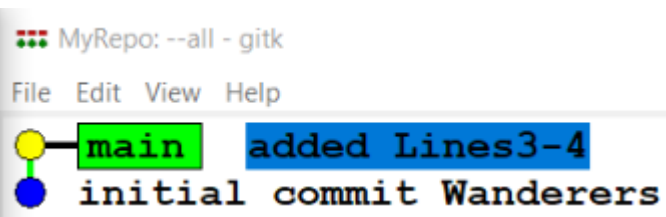
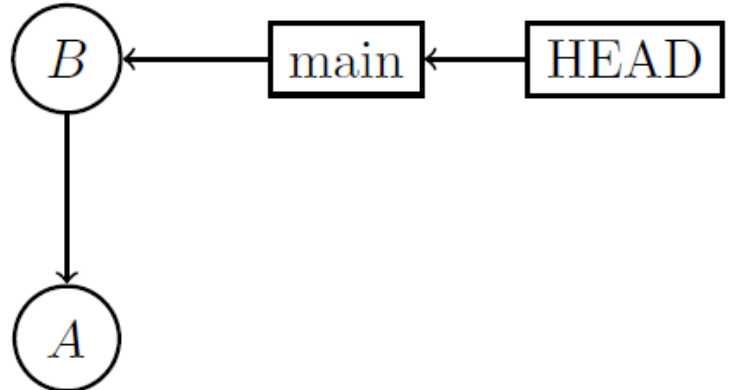
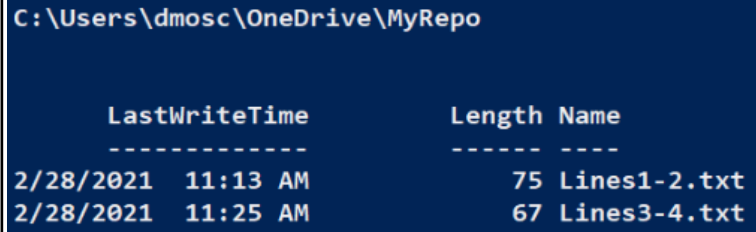
A Git repository is a digraph

Vertex / Commit, Pointer

Edge / parent/child

Path / branch



<div><div>Git language</div><div>Make a second commit to the main branch.</div></div>	<div><div>Digraph language</div><div>Create a new vertex, B, with parent A. Change the referent of pointer main from A to B.</div></div>	<div><div>Command line</div><div>Git commit -m "added Lines3-4"</div></div>
<div><div>gitk before change</div><div>A screenshot of the gitk GUI. The title bar says 'MyRepo: --all - gitk'. The menu bar has 'File', 'Edit', 'View', and 'Help'. The main area shows a yellow circle representing the initial commit 'Wanderers' with a green box labeled 'main' pointing to it. The text 'initial commit Wanderers' is displayed.</div></div>	<div><div>Digraph before change</div><div>A digraph showing a linear history. A circle labeled 'A' has an arrow pointing to a box labeled 'main', which in turn has an arrow pointing to a box labeled 'HEAD'.</div></div>	<div><div>Working directory before change</div><div>A terminal window showing the contents of the working directory. The path is 'C:\Users\dmosc\OneDrive\MyRepo'. It lists two files: 'Lines1-2.txt' (75 bytes, last written 2/28/2021 11:13 AM) and 'Lines3-4.txt' (67 bytes, last written 2/28/2021 11:25 AM).</div></div>
<div><div>gitk after change</div><div>A screenshot of the gitk GUI after a commit. The title bar is the same. The main area shows the yellow circle 'Wanderers' with 'main' pointing to it. A new blue circle is shown below 'Wanderers', and the text 'added Lines3-4' is displayed.</div></div>	<div><div>Digraph after change</div><div>A digraph showing a new commit. A circle labeled 'B' has an arrow pointing to a box labeled 'main', which points to a box labeled 'HEAD'. A vertical arrow points from 'B' down to a circle labeled 'A'.</div></div>	<div><div>Working directory after change</div><div>A terminal window showing the same file listing as before, but now it includes 'Lines3-4.txt' (67 bytes, last written 2/28/2021 11:25 AM) alongside 'Lines1-2.txt'.</div></div>

Git language

Make a second commit to the main branch.

gitk before change

```
MyRepo: --all - gitk
File Edit View Help
● main initial commit Wanderers
```

gitk after change

```
MyRepo: --all - gitk
File Edit View Help
● main added Lines3-4
● initial commit Wanderers
```

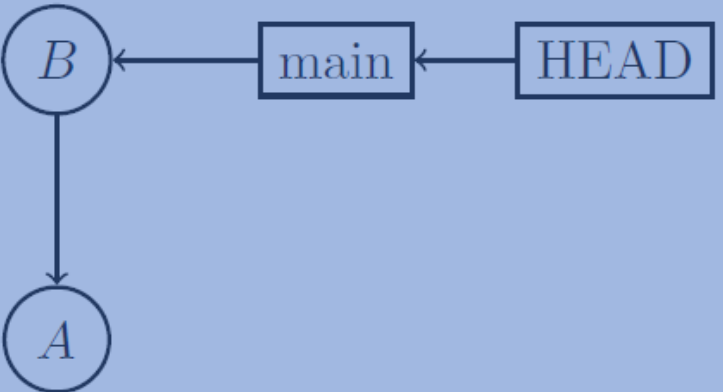
Digraph language

*Create a new vertex, B, with parent A.
Change the referent of pointer main from A to B.*

Digraph before change



Digraph after change



Command line

`Git commit -m "added Lines3-4"`

Working directory before change

C:\Users\dmosc\OneDrive\MyRepo

	LastWriteTime	Length	Name
	-----	-----	----
2/28/2021	11:13 AM	75	Lines1-2.txt
2/28/2021	11:25 AM	67	Lines3-4.txt

Working directory after change

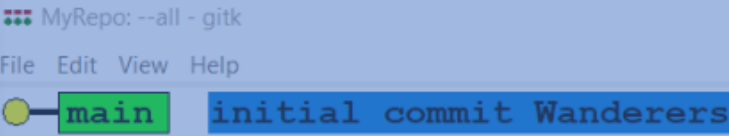
C:\Users\dmosc\OneDrive\MyRepo

	LastWriteTime	Length	Name
	-----	-----	----
2/28/2021	11:13 AM	75	Lines1-2.txt
2/28/2021	11:25 AM	67	Lines3-4.txt

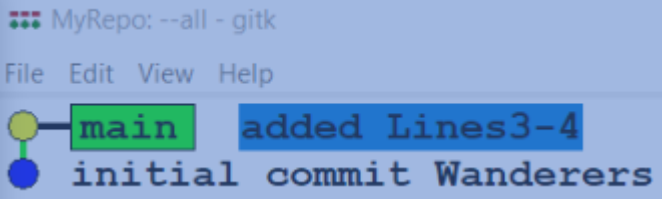
Git language

Make a second commit to the main branch.

gitk before change



gitk after change



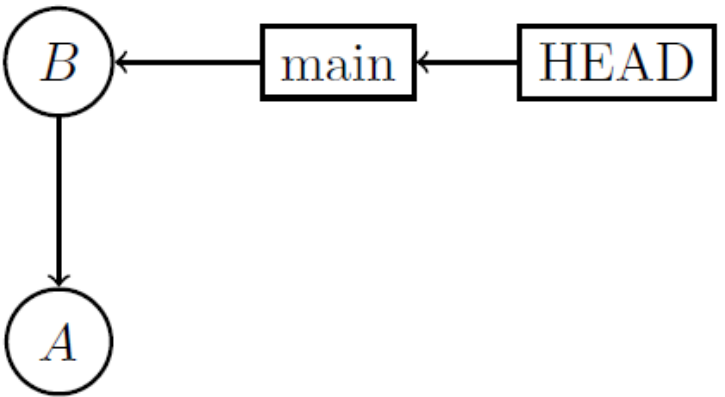
Digraph language

Create a new vertex, B, with parent A.
Change the referent of pointer main from A to B.

Digraph before change



Digraph after change



Command line

```
Git commit -m "added Lines3-4"
```

Working directory before change

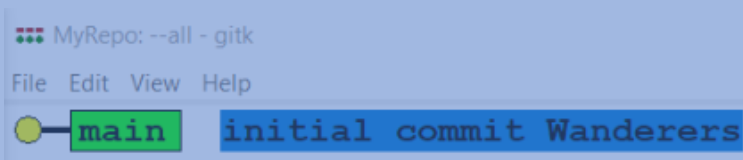

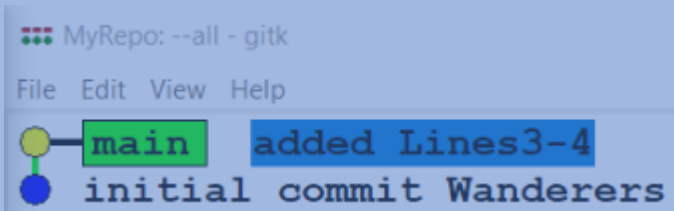
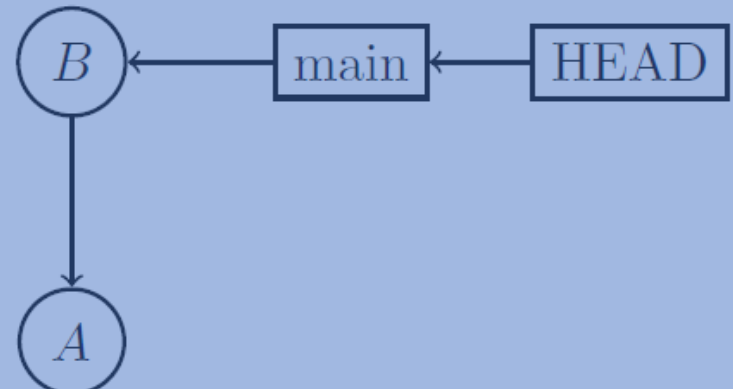
C:\Users\dmosc\OneDrive\MyRepo

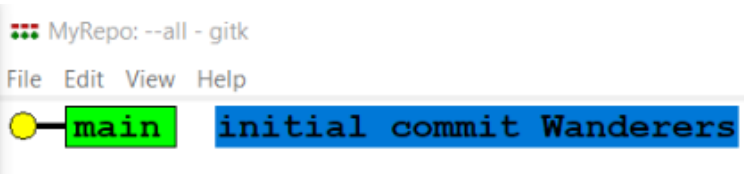

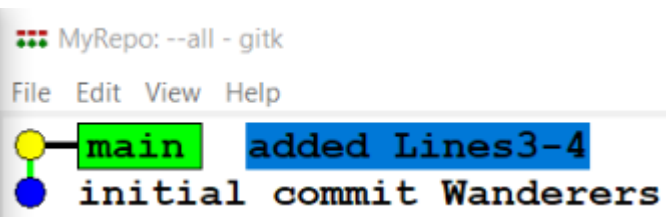
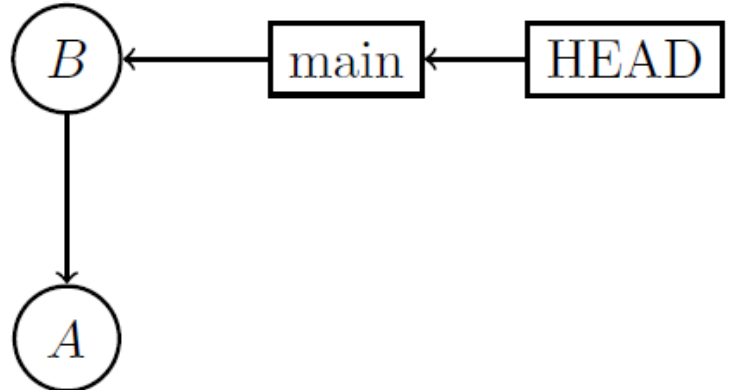
	LastWriteTime	Length	Name
	-----	-----	----
	2/28/2021 11:13 AM	75	Lines1-2.txt
	2/28/2021 11:25 AM	67	Lines3-4.txt

Working directory after change

C:\Users\dmosc\OneDrive\MyRepo

	LastWriteTime	Length	Name
	-----	-----	----
	2/28/2021 11:13 AM	75	Lines1-2.txt
	2/28/2021 11:25 AM	67	Lines3-4.txt

<div>Git language</div> <div>Make a second commit to the main branch.</div>	<div>Digraph language</div> <div>Create a new vertex, B, with parent A. Change the referent of pointer main from A to B.</div>	<div>Command line</div> <div>Git commit -m "added Lines3-4"</div>																
<div>gitk before change</div> <div></div>	<div>Digraph before change</div> <div></div>	<div>Working directory before change</div> <div><pre>C:\Users\dmosc\OneDrive\MyRepo</pre><table><thead><tr><th></th><th>LastWriteTime</th><th>Length</th><th>Name</th></tr><tr><th></th><th>-----</th><th>-----</th><th>----</th></tr></thead><tbody><tr><td>2/28/2021</td><td>11:13 AM</td><td>75</td><td>Lines1-2.txt</td></tr><tr><td>2/28/2021</td><td>11:25 AM</td><td>67</td><td>Lines3-4.txt</td></tr></tbody></table></div>		LastWriteTime	Length	Name		-----	-----	----	2/28/2021	11:13 AM	75	Lines1-2.txt	2/28/2021	11:25 AM	67	Lines3-4.txt
	LastWriteTime	Length	Name															
	-----	-----	----															
2/28/2021	11:13 AM	75	Lines1-2.txt															
2/28/2021	11:25 AM	67	Lines3-4.txt															
<div>gitk after change</div> <div></div>	<div>Digraph after change</div> <div></div>	<div>Working directory after change</div> <div><pre>C:\Users\dmosc\OneDrive\MyRepo</pre><table><thead><tr><th></th><th>LastWriteTime</th><th>Length</th><th>Name</th></tr><tr><th></th><th>-----</th><th>-----</th><th>----</th></tr></thead><tbody><tr><td>2/28/2021</td><td>11:13 AM</td><td>75</td><td>Lines1-2.txt</td></tr><tr><td>2/28/2021</td><td>11:25 AM</td><td>67</td><td>Lines3-4.txt</td></tr></tbody></table></div>		LastWriteTime	Length	Name		-----	-----	----	2/28/2021	11:13 AM	75	Lines1-2.txt	2/28/2021	11:25 AM	67	Lines3-4.txt
	LastWriteTime	Length	Name															
	-----	-----	----															
2/28/2021	11:13 AM	75	Lines1-2.txt															
2/28/2021	11:25 AM	67	Lines3-4.txt															

Git language <i>Make a second commit to the main branch.</i>	Digraph language <i>Create a new vertex, B, with parent A. Change the referent of pointer main from A to B.</i>	Command line <pre>Git commit -m "added Lines3-4"</pre>									
gitk before change 	Digraph before change 	Working directory before change <pre>C:\Users\dmosc\OneDrive\MyRepo</pre> <table border="1"> <thead> <tr> <th>LastWriteTime</th> <th>Length</th> <th>Name</th> </tr> </thead> <tbody> <tr> <td>2/28/2021 11:13 AM</td> <td>75</td> <td>Lines1-2.txt</td> </tr> <tr> <td>2/28/2021 11:25 AM</td> <td>67</td> <td>Lines3-4.txt</td> </tr> </tbody> </table>	LastWriteTime	Length	Name	2/28/2021 11:13 AM	75	Lines1-2.txt	2/28/2021 11:25 AM	67	Lines3-4.txt
LastWriteTime	Length	Name									
2/28/2021 11:13 AM	75	Lines1-2.txt									
2/28/2021 11:25 AM	67	Lines3-4.txt									
gitk after change 	Digraph after change 	Working directory after change <pre>C:\Users\dmosc\OneDrive\MyRepo</pre> <table border="1"> <thead> <tr> <th>LastWriteTime</th> <th>Length</th> <th>Name</th> </tr> </thead> <tbody> <tr> <td>2/28/2021 11:13 AM</td> <td>75</td> <td>Lines1-2.txt</td> </tr> <tr> <td>2/28/2021 11:25 AM</td> <td>67</td> <td>Lines3-4.txt</td> </tr> </tbody> </table>	LastWriteTime	Length	Name	2/28/2021 11:13 AM	75	Lines1-2.txt	2/28/2021 11:25 AM	67	Lines3-4.txt
LastWriteTime	Length	Name									
2/28/2021 11:13 AM	75	Lines1-2.txt									
2/28/2021 11:25 AM	67	Lines3-4.txt									

Summary



Git commands can often be understood as modifications of a directed graph.

Working with a directed graph model is one way to gain conceptual understanding of a Git repository.

Next steps

- Expand the template to include other representations: e.g., remote and local repos.
- Use multiple representations to clearly state / plan solutions for problems.
- Quiz yourself!
 - Choose two distinct valid digraphs. What Git/Hub commands transform one into the other?



References and resources

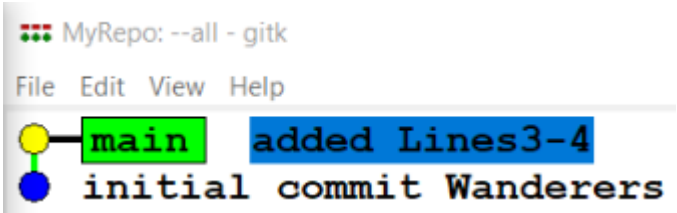
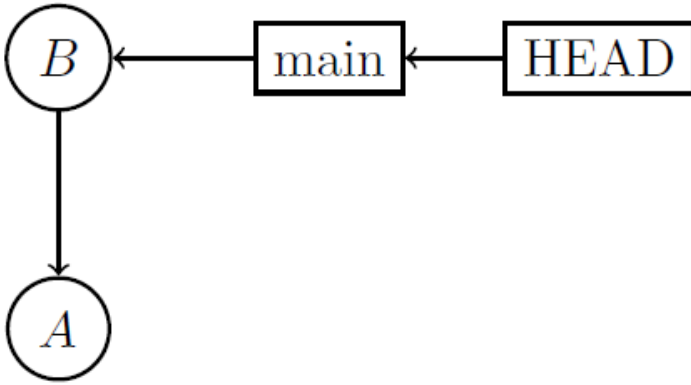
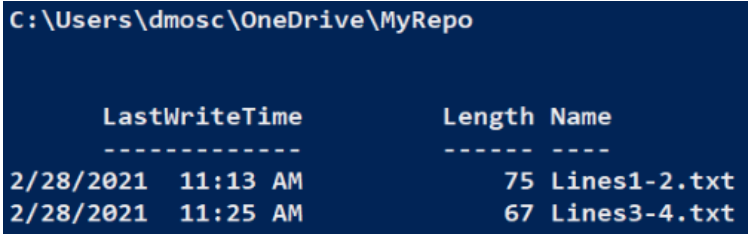
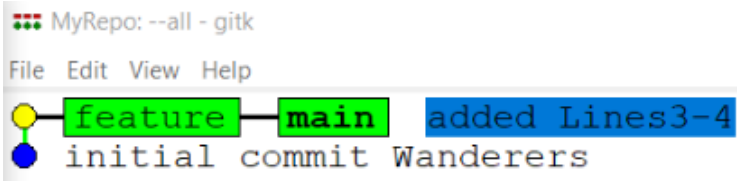
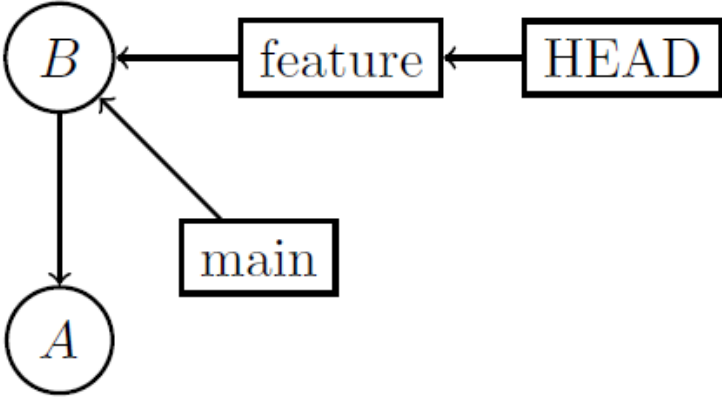
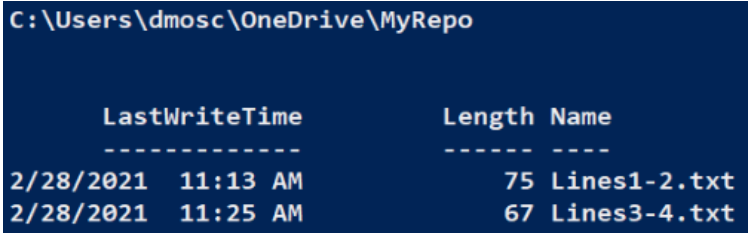
*O! Wanderers in the shadowed land
despair not! For though dark they stand,
all woods there be must end at last,
and see the open sun go past:
the setting sun, the rising sun,
the day's end, or the day begun.
For east or west all woods must fail...*

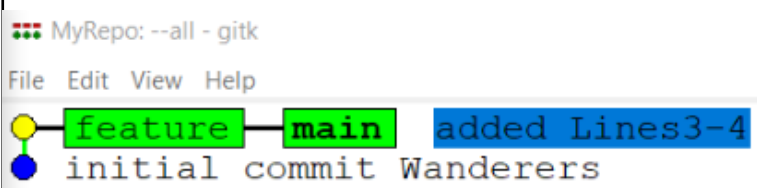
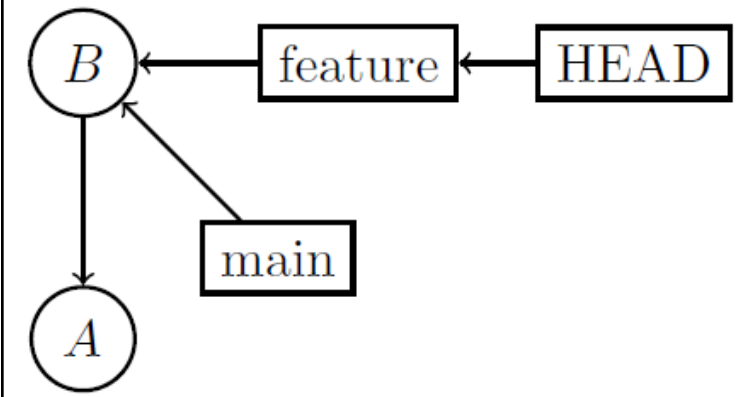
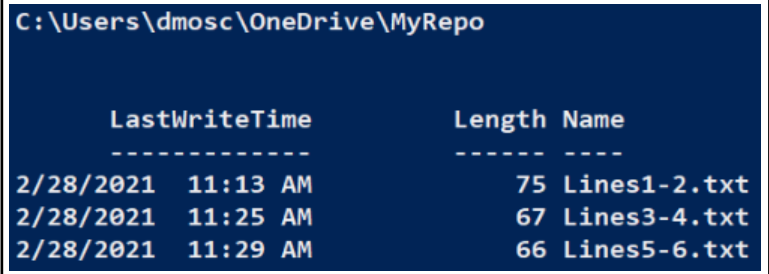
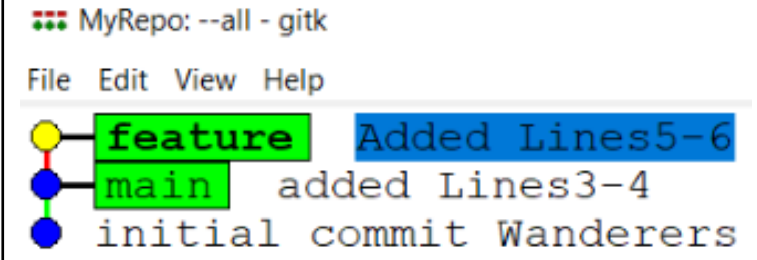
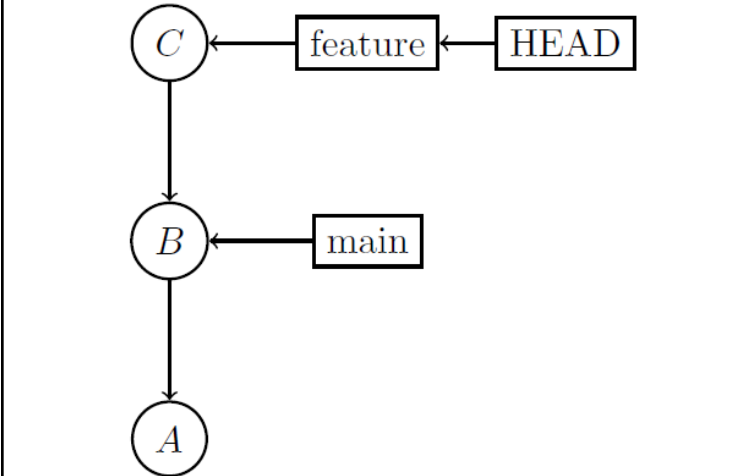
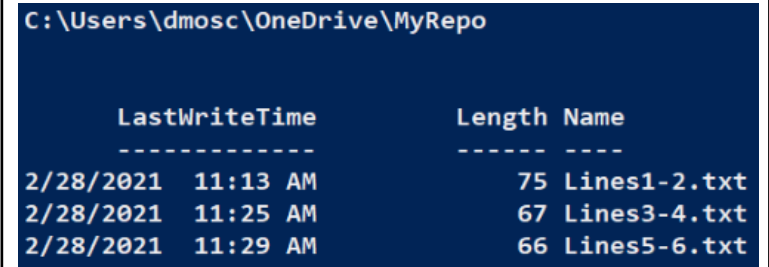
Chacon and Straub. *Pro Git*, Second Edition. Version 2.1.236, 2020-07-06.

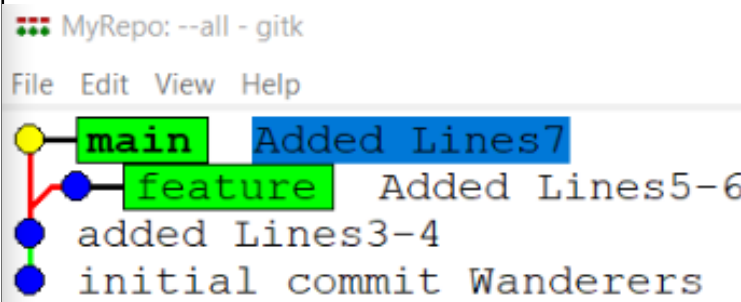
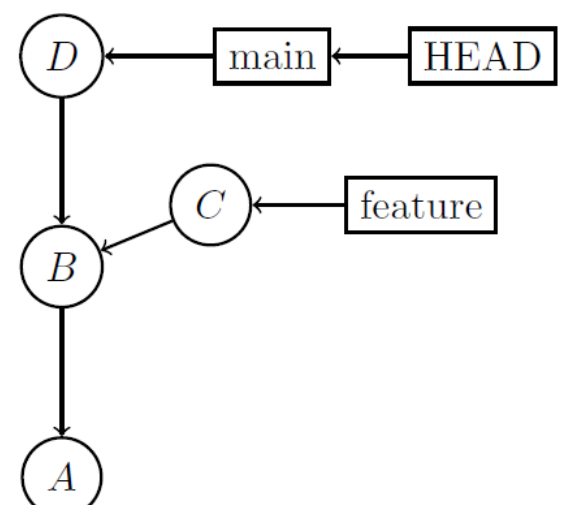
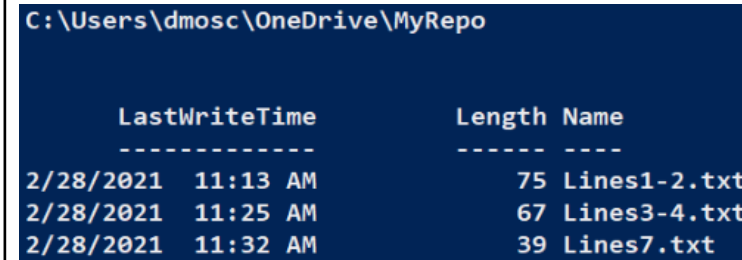
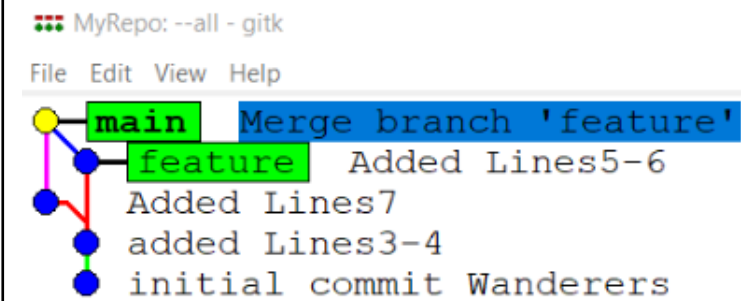
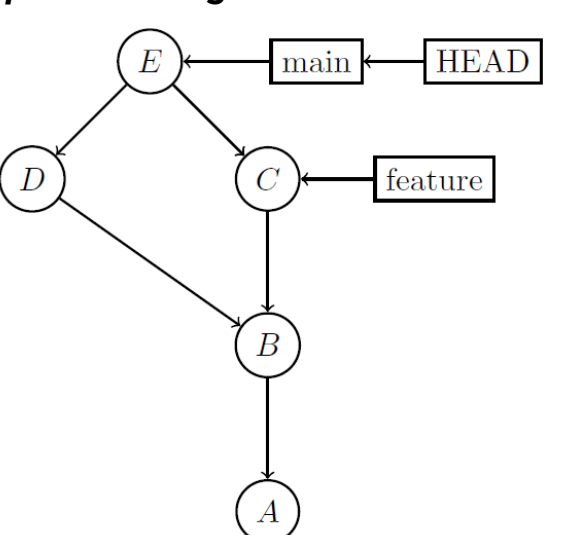
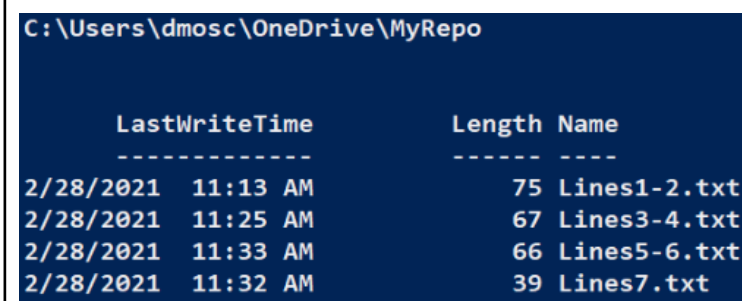
Livingston-Gray, Sam. think-like-a-git.net.

Tomayko, Ryan. "The Thing About Git."

<https://tomayko.com/blog/2008/the-thing-about-git>

Git language	Digraph language	Command line									
<p>Create the branch feature. Checkout the new branch.</p>	<p>Create a new pointer, feature, with referent B. Change the referent of HEAD from main to feature.</p>	<pre>git branch feature git checkout feature</pre>									
<p>gitk before change</p>  <p>A screenshot of the gitk GUI. The title bar says 'MyRepo: --all - gitk'. The menu bar has 'File', 'Edit', 'View', and 'Help'. The main area shows a commit graph with a yellow circle (HEAD) pointing to a green box labeled 'main'. Below the graph, it says 'added Lines3-4' and 'initial commit Wanderers'.</p>	<p>Digraph before change</p>  <p>A digraph showing the state before the change. It has three nodes: a circle labeled 'B', a rectangle labeled 'main', and a circle labeled 'A'. Arrows point from 'HEAD' (rectangle) to 'main' (rectangle), and from 'main' (rectangle) to 'B' (circle). An arrow points from 'B' (circle) to 'A' (circle).</p>	<p>Working directory before change</p>  <p>A screenshot of a terminal window showing the contents of the working directory. The title bar says 'C:\Users\dmosc\OneDrive\MyRepo'. The output shows a list of files with their last write time, length, and name.</p> <table border="1"> <thead> <tr> <th>LastWriteTime</th> <th>Length</th> <th>Name</th> </tr> </thead> <tbody> <tr> <td>2/28/2021 11:13 AM</td> <td>75</td> <td>Lines1-2.txt</td> </tr> <tr> <td>2/28/2021 11:25 AM</td> <td>67</td> <td>Lines3-4.txt</td> </tr> </tbody> </table>	LastWriteTime	Length	Name	2/28/2021 11:13 AM	75	Lines1-2.txt	2/28/2021 11:25 AM	67	Lines3-4.txt
LastWriteTime	Length	Name									
2/28/2021 11:13 AM	75	Lines1-2.txt									
2/28/2021 11:25 AM	67	Lines3-4.txt									
<p>gitk after change</p>  <p>A screenshot of the gitk GUI after the change. The title bar says 'MyRepo: --all - gitk'. The menu bar has 'File', 'Edit', 'View', and 'Help'. The main area shows a commit graph with a yellow circle (HEAD) pointing to a green box labeled 'feature'. A green box labeled 'main' is also shown, with an arrow pointing from 'main' to 'feature'. Below the graph, it says 'added Lines3-4' and 'initial commit Wanderers'.</p>	<p>Digraph after change</p>  <p>A digraph showing the state after the change. It has four nodes: a circle labeled 'B', a rectangle labeled 'feature', a rectangle labeled 'main', and a circle labeled 'A'. Arrows point from 'HEAD' (rectangle) to 'feature' (rectangle), and from 'feature' (rectangle) to 'B' (circle). An arrow points from 'main' (rectangle) to 'feature' (rectangle). An arrow points from 'B' (circle) to 'A' (circle).</p>	<p>Working directory after change</p>  <p>A screenshot of a terminal window showing the contents of the working directory after the change. The title bar says 'C:\Users\dmosc\OneDrive\MyRepo'. The output shows a list of files with their last write time, length, and name.</p> <table border="1"> <thead> <tr> <th>LastWriteTime</th> <th>Length</th> <th>Name</th> </tr> </thead> <tbody> <tr> <td>2/28/2021 11:13 AM</td> <td>75</td> <td>Lines1-2.txt</td> </tr> <tr> <td>2/28/2021 11:25 AM</td> <td>67</td> <td>Lines3-4.txt</td> </tr> </tbody> </table>	LastWriteTime	Length	Name	2/28/2021 11:13 AM	75	Lines1-2.txt	2/28/2021 11:25 AM	67	Lines3-4.txt
LastWriteTime	Length	Name									
2/28/2021 11:13 AM	75	Lines1-2.txt									
2/28/2021 11:25 AM	67	Lines3-4.txt									

Git language <i>Add a commit to the branch feature.</i>	Digraph language <i>Create vertex C with parent B. Change the referent of pointer feature to C.</i>	Command line git commit "added Lines5-6"
gitk before change  <p>MyRepo: --all - gitk File Edit View Help feature main added Lines3-4 initial commit Wanderers</p>	Digraph before change  <pre> graph TD B((B)) --> A((A)) B --> main1[main] feature[feature] --> B HEAD[HEAD] --> feature </pre>	Working directory before change  <pre> C:\Users\dmosc\OneDrive\MyRepo LastWriteTime Length Name - 2/28/2021 11:13 AM 75 Lines1-2.txt 2/28/2021 11:25 AM 67 Lines3-4.txt 2/28/2021 11:29 AM 66 Lines5-6.txt </pre>
gitk after change  <p>MyRepo: --all - gitk File Edit View Help feature Added Lines5-6 main added Lines3-4 initial commit Wanderers</p>	Digraph after change  <pre> graph TD C((C)) --> B((B)) C --> feature[feature] B --> A((A)) B --> main[main] feature --> C HEAD[HEAD] --> feature </pre>	Working directory after change  <pre> C:\Users\dmosc\OneDrive\MyRepo LastWriteTime Length Name - 2/28/2021 11:13 AM 75 Lines1-2.txt 2/28/2021 11:25 AM 67 Lines3-4.txt 2/28/2021 11:29 AM 66 Lines5-6.txt </pre>

<div>Git language</div> <div>Merge feature branch to main branch</div>	<div>Digraph language</div> <div>Create vertex E with parents C and D. Change referent of pointer main to E.</div>	<div>Command line</div> <div>git checkout main git merge feature</div>
<div>gitk before change</div> <div></div>	<div>Digraph before change</div> <div></div>	<div>Working directory before change</div> <div></div>
<div>gitk after change</div> <div></div>	<div>Digraph after change</div> <div></div>	<div>Working directory after change</div> <div></div>

Git language

Enter detached head state by checking out initial commit. Commit while in detached head state.

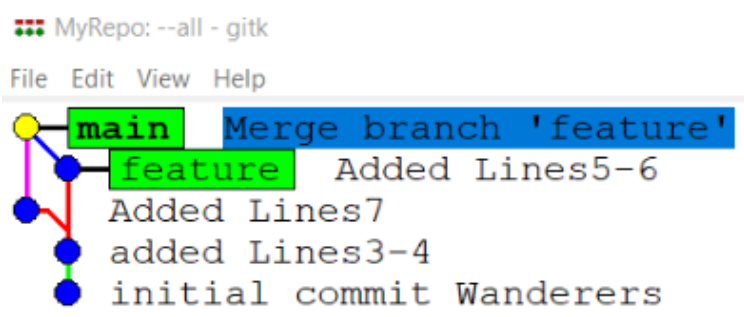
Digraph language

Change referent of HEAD to A. Create vertex F with parent A. Change referent of HEAD to F.

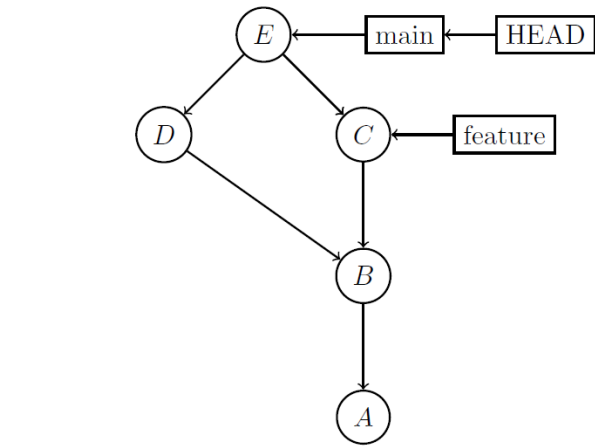
Command line

Git checkout 1b8348e1a
[;]
Git commit -m "added whole poem while in detached head state"

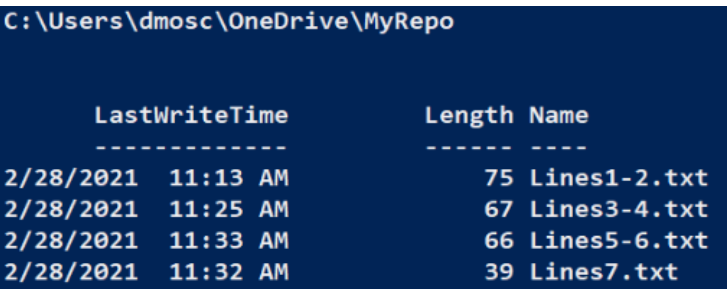
gitk before change



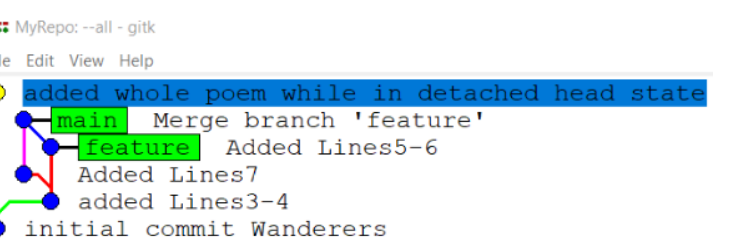
Digraph before change



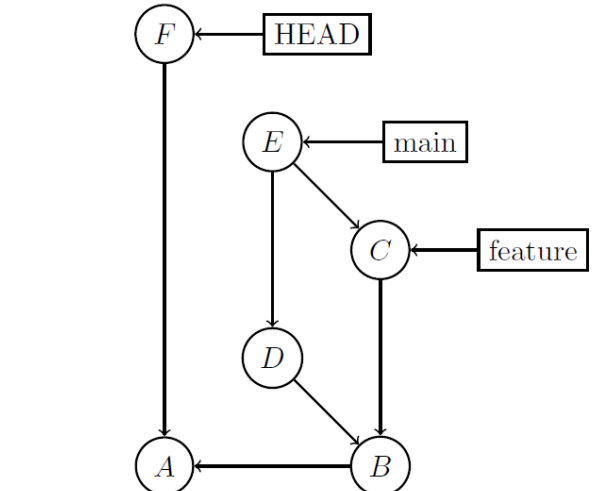
Working directory before change



gitk after change



Digraph after change



Working directory after change

