TEMPO Workshop on Software Development

Michal Kvasnica



SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA "An expert is a person who has made all the mistakes that can be made in a very narrow field."

Niels Bohr, Nobel price winner (1922)

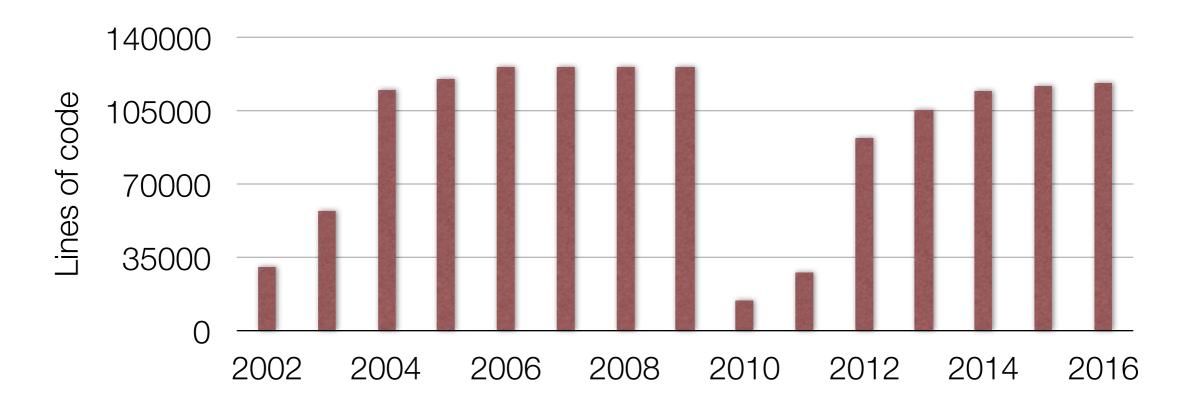
The Multi-Parametric Toolbox (MPT)

Developed since 2002

40,000+ installations world-wide

118,503 lines of code in 2,169 files

5 main developers



Agenda

Today

- version control systems and collaborative development
- Mercurial, Git, Bitbucket, GitHub
- providing support

Tomorrow

- unit testing
- documentation
- dissemination

Set of DOs and DONTs

Agenda

Today

- version control systems and collaborative development

- Mercurial, Git, Bitbucket, GitHub
- providing support

Tomorrow

- unit testing
- documentation
- dissemination

DONT #1: Do NOT Use Archives

mpt12.zip mpt1 2 2.zip mpt ver04.zip mpt ver11R1.zip mpt ver141.zip mpt13r1.zip mpt ver02.zip mpt ver05.zip mpt ver11R2.zip mpt ver144.zip mpt14.zip mpt ver03.zip mpt ver11.zip mpt ver11R3.zip mptv122.zip

Main problems:

- does not protect against accidental deletion
- linear development (cannot release until all new features are complete)
- single-user development
- difficult to find a change that introduced a bug

Archives do not give answers to:

- what has changed between versions?
- who made the changes?
- when and why were the changes made?
- which version is the latest stable release?

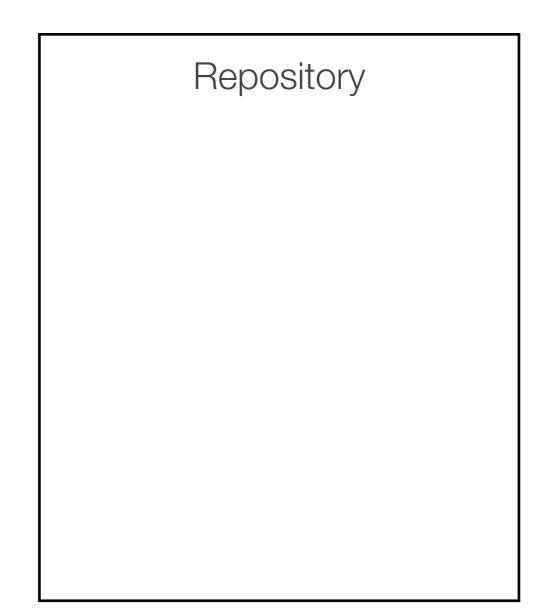
- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

Main objectives:

- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

Working directory

file1 file2



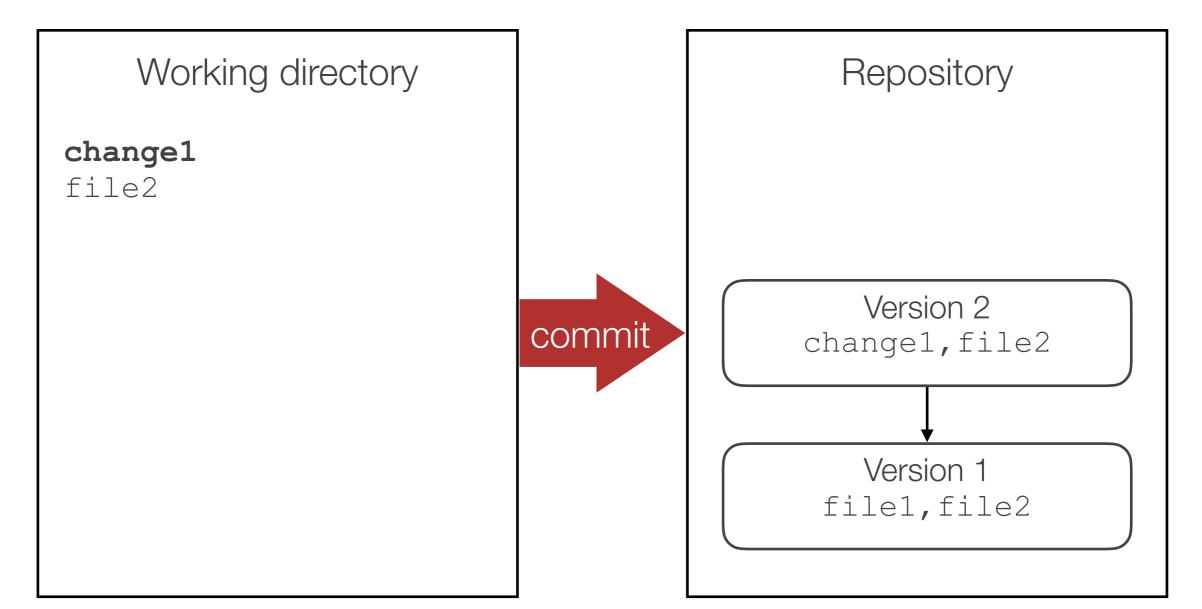
- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

Working directory		Repository
file1 file2		
	commit	Version 1 file1, file2

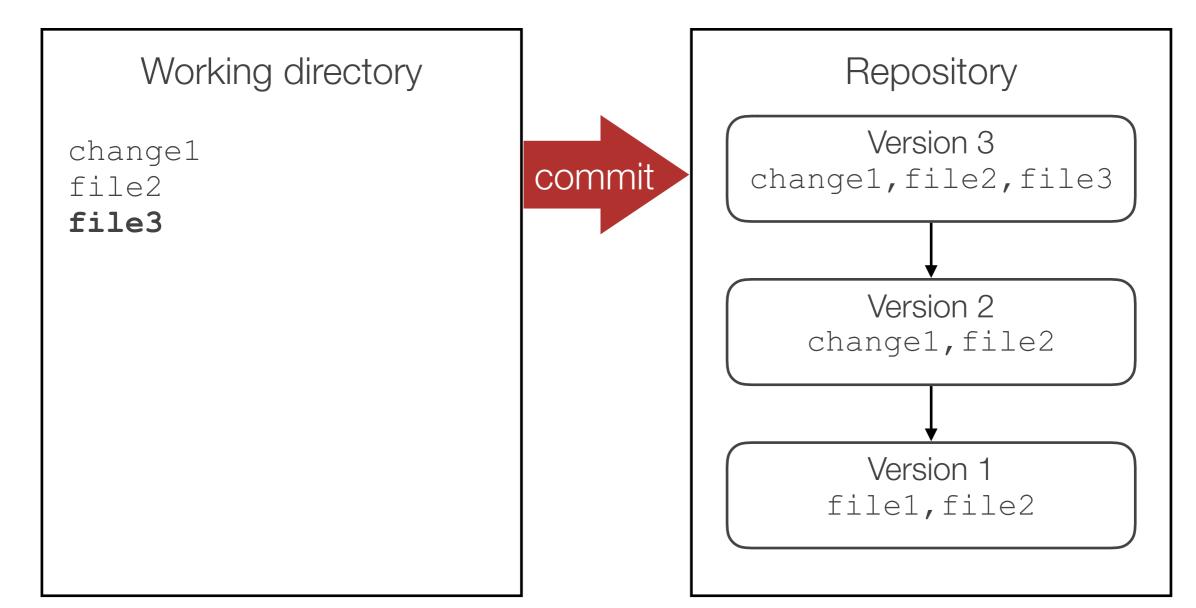
Main objectives:

- record changes over time

- recall a specific version later
- enable collaboration
- allow nonlinear development



- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development



Sidenote: VCS is for Source Code

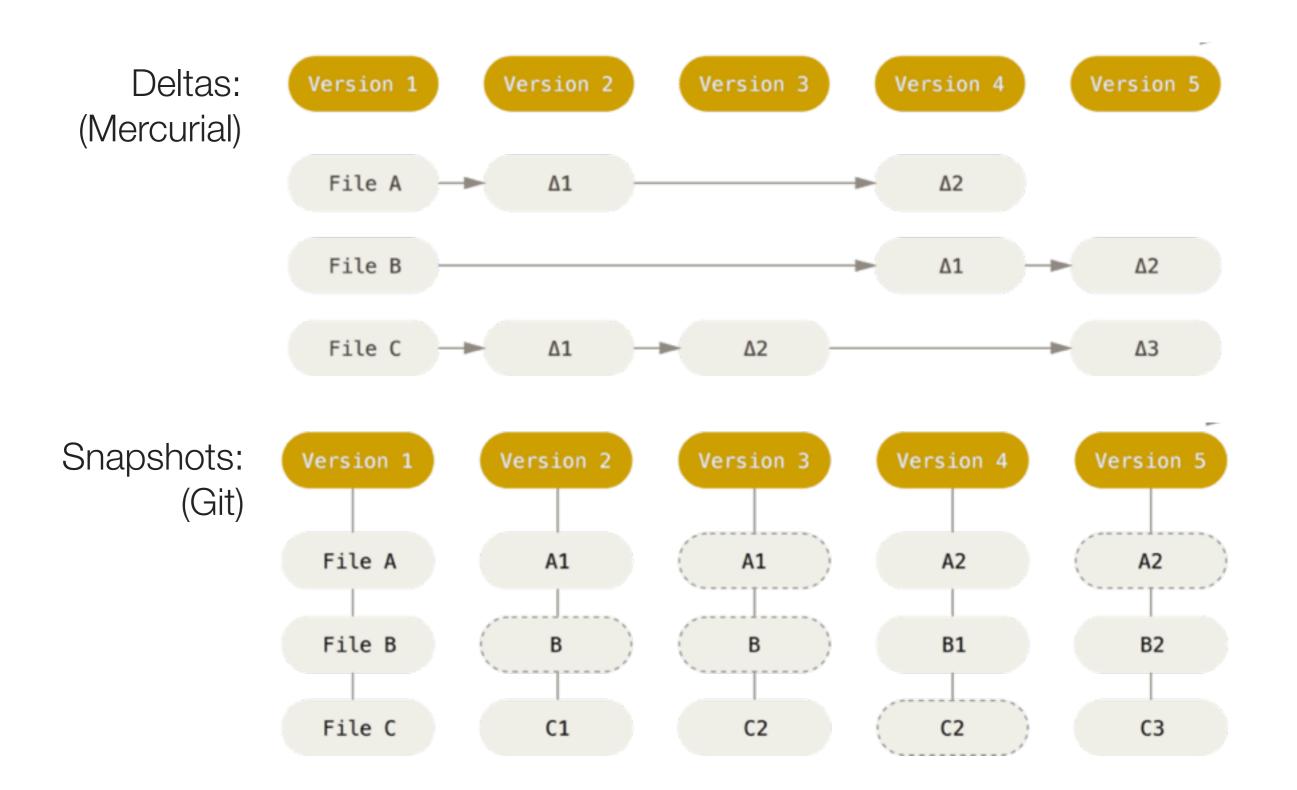
Typically stored in VCS:

- source files
- documentation (markdown, html, LaTeX)
- binary files that do not change (e.g. logos)

Not to be stored:

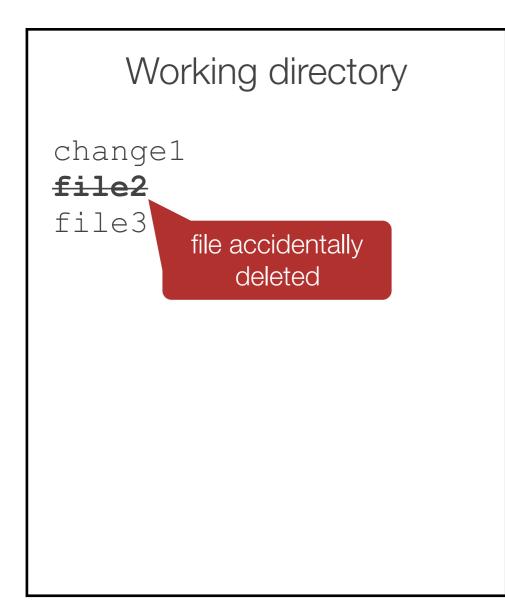
- automatically generated files (PDFs from LaTeX, compiled binaries, etc.)
- Word documents
- files containing passwords! (even when you remove the data from the top version, it stays in the repository forever in previous versions)

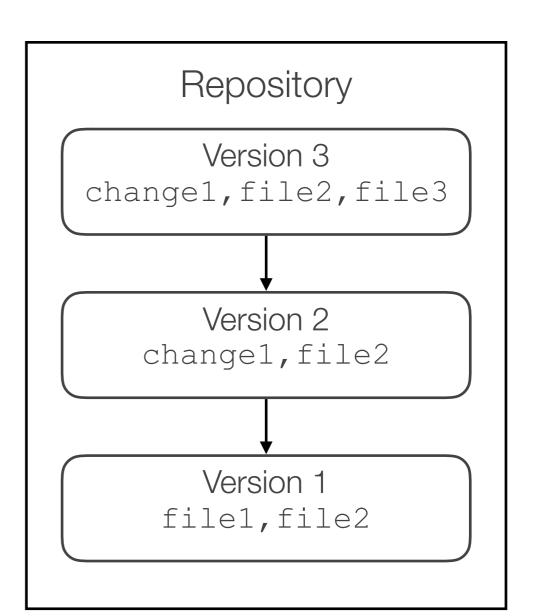
Sidenote: Different Storage Models



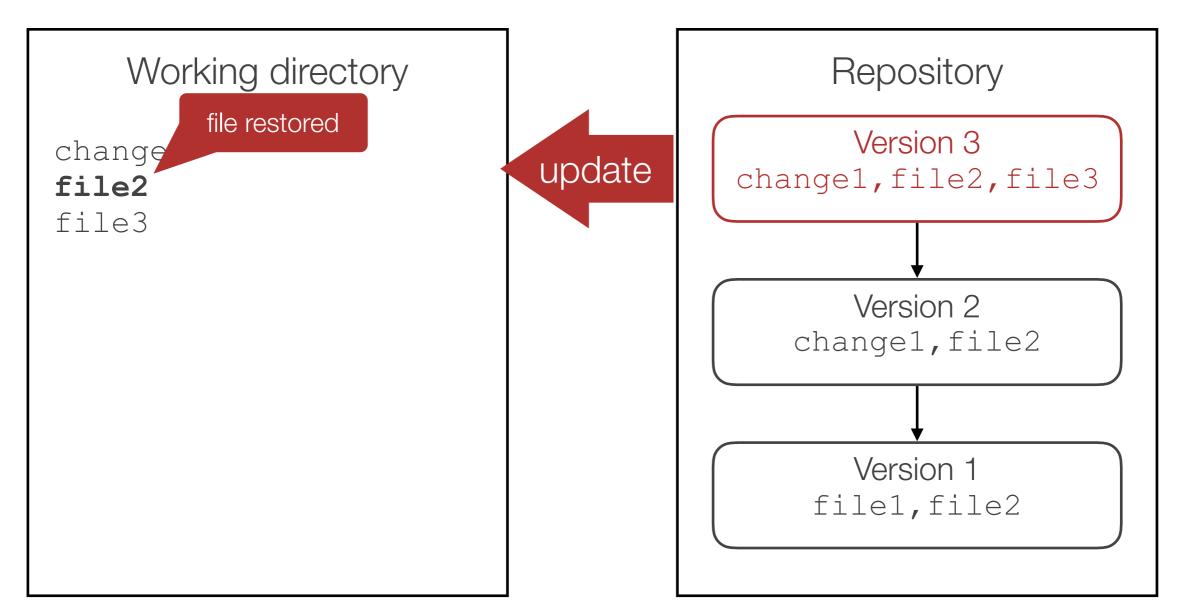
https://git-scm.com/book/en/v2/Getting-Started-Git-Basics

- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

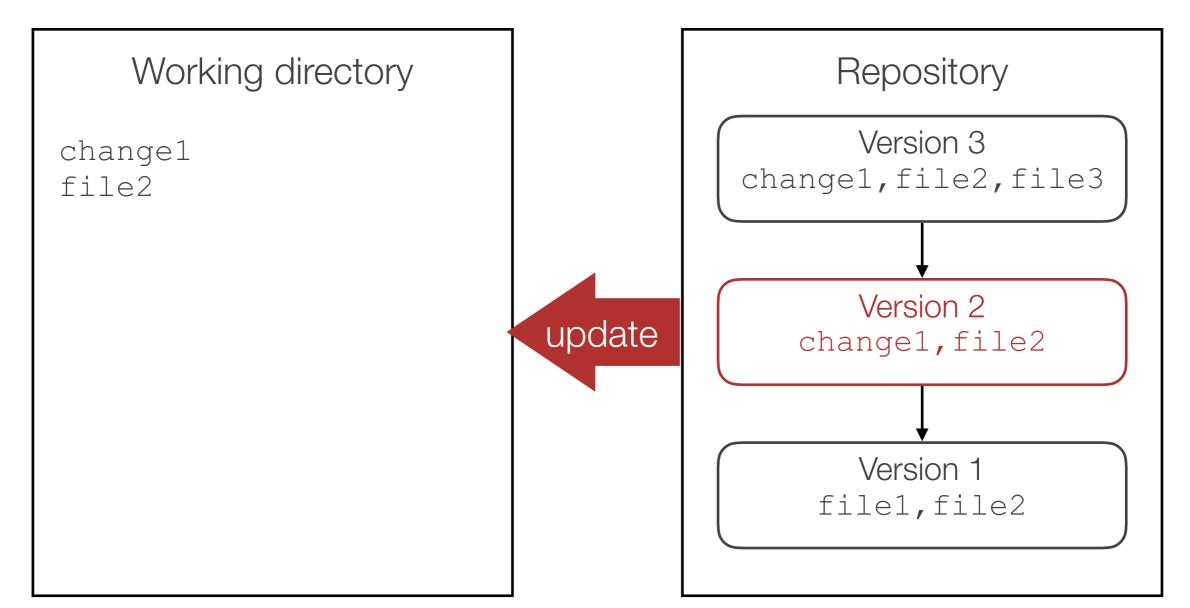




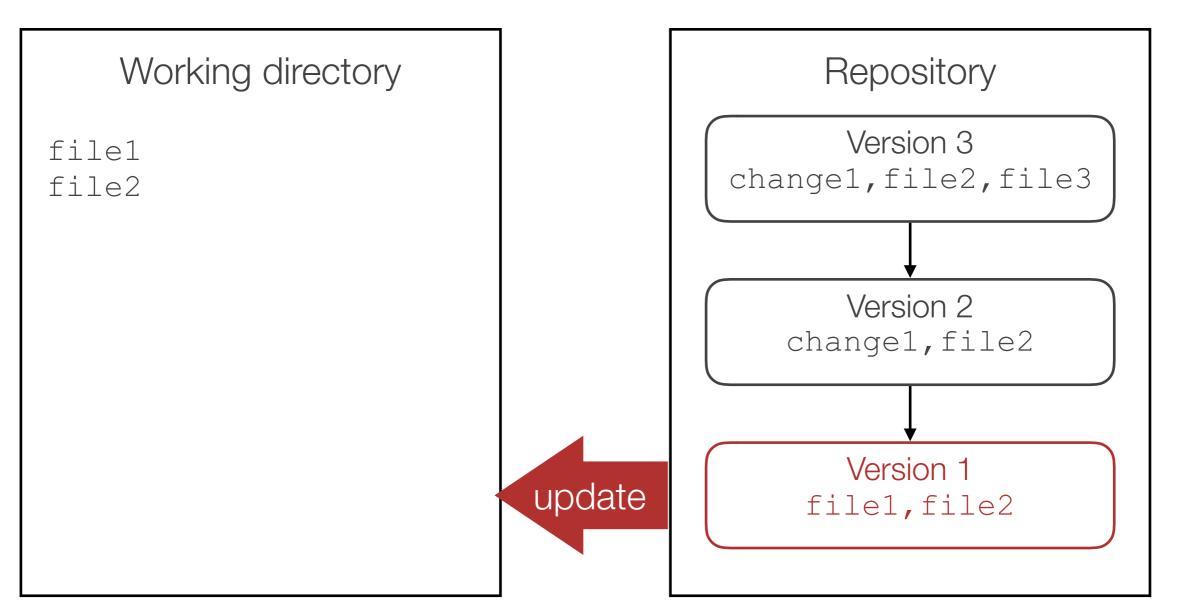
- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development



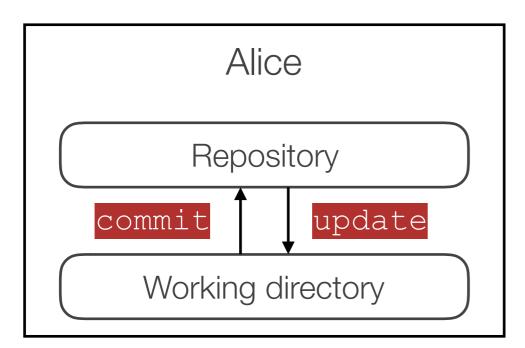
- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

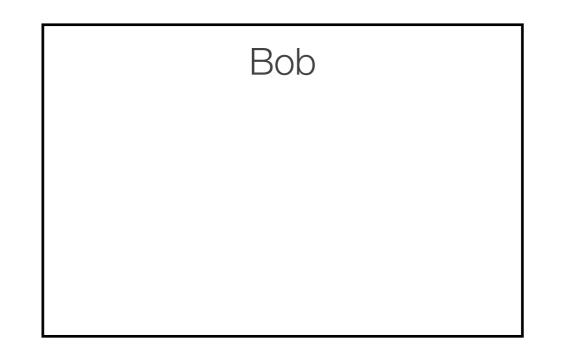


- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

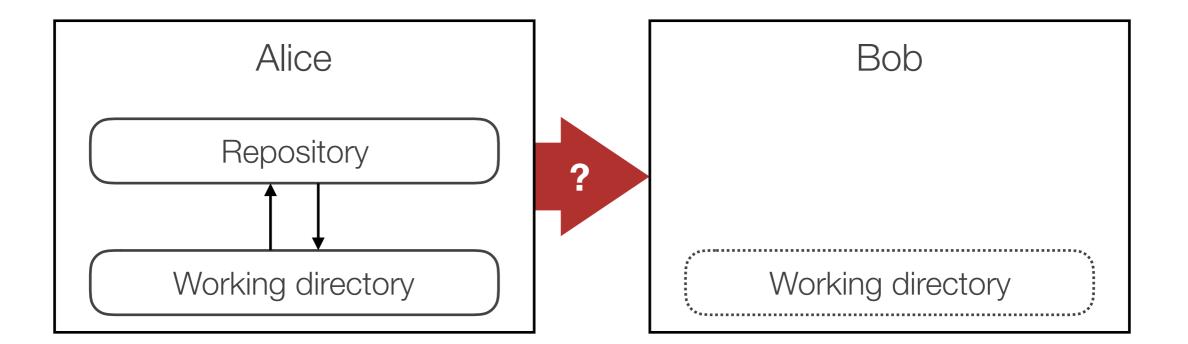


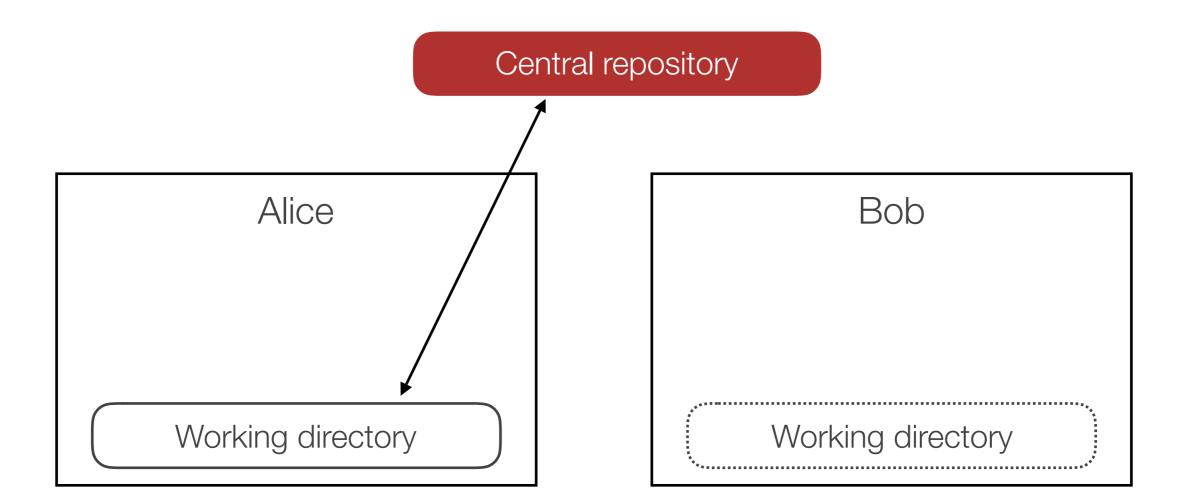
- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

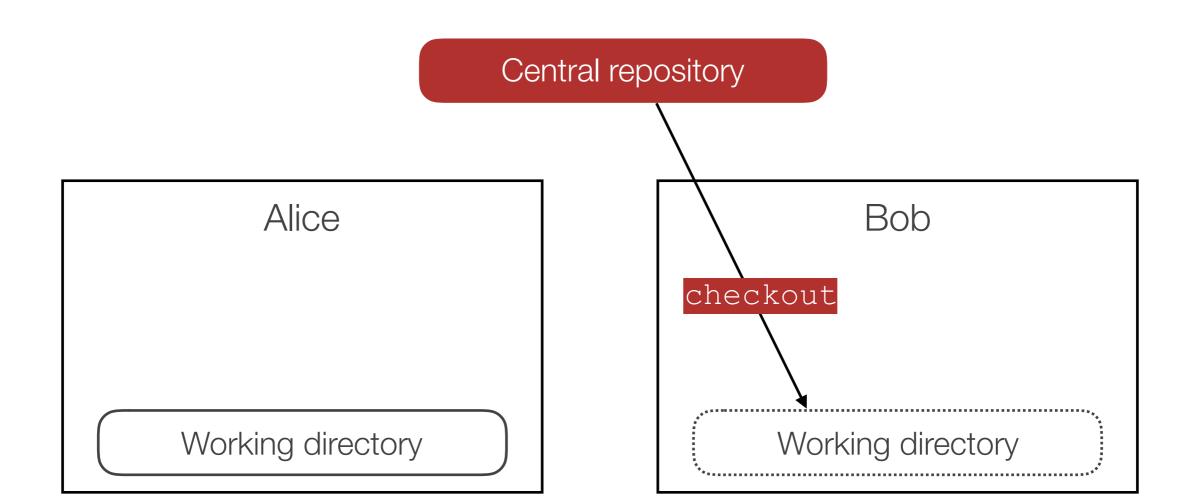


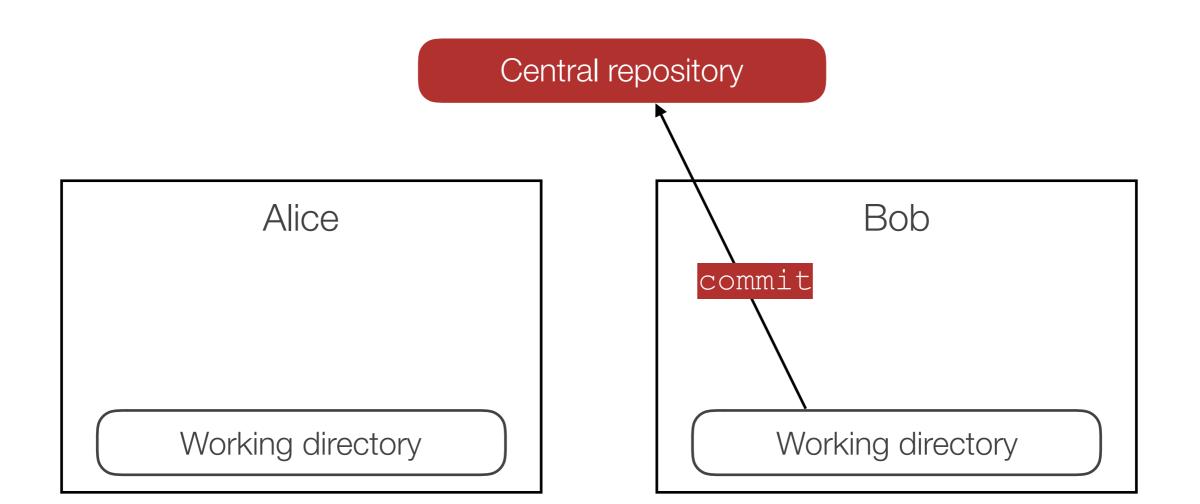


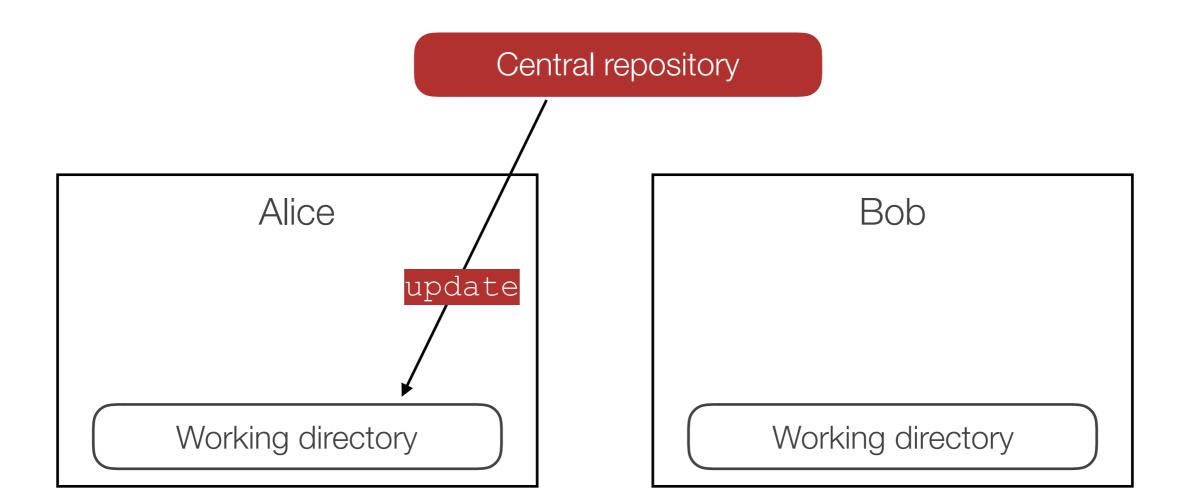
- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

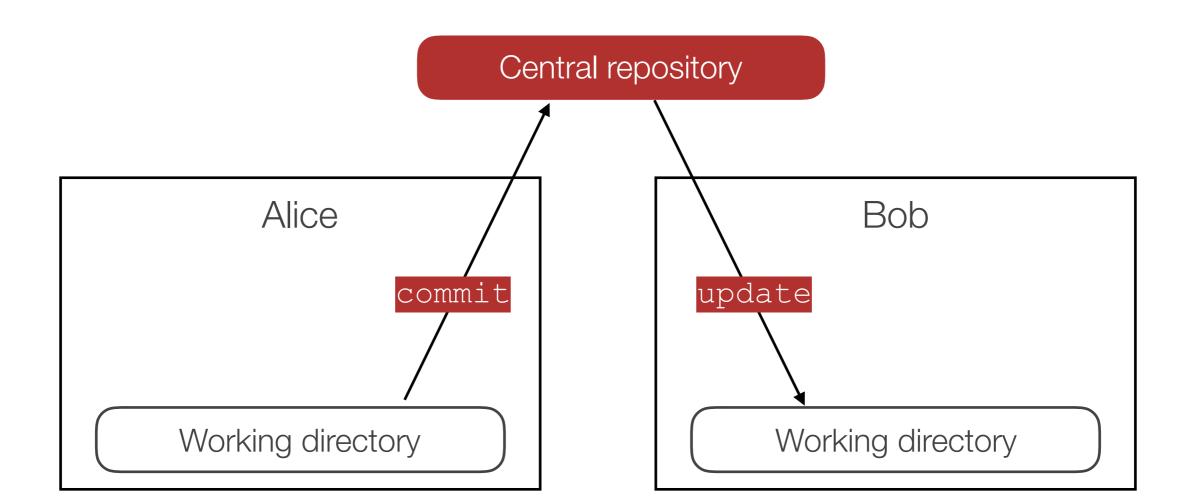








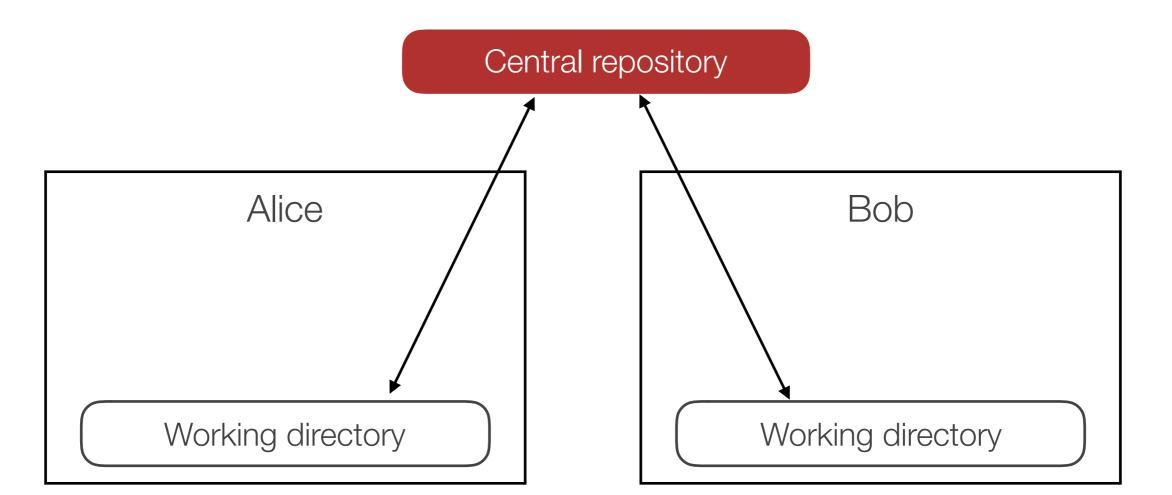




Main idea: use a centralized repository (public server)

Problems:

- requires permanent internet connection (+single point of failure)
- access rights management
- communication ("Hey, I made a change, you need to update")
- only one person can commit at a time

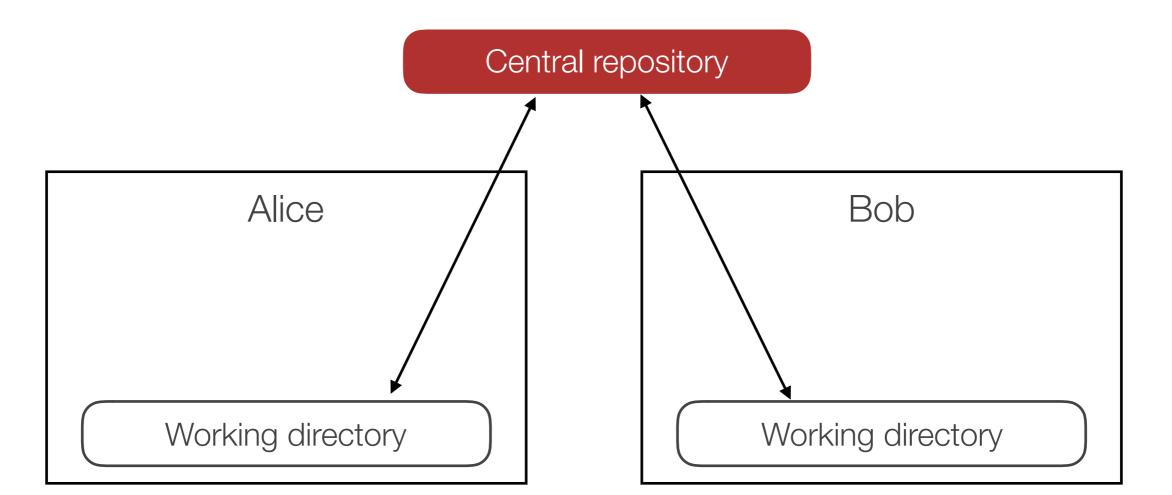


DONT #2: Never use a Centralized VCS

Main idea: use a centralized repository (public server)

Problems:

- requires permanent internet connection (+single point of failure)
- access rights management
- communication ("Hey, I made a change, you need to update")
- only one person can commit at a time

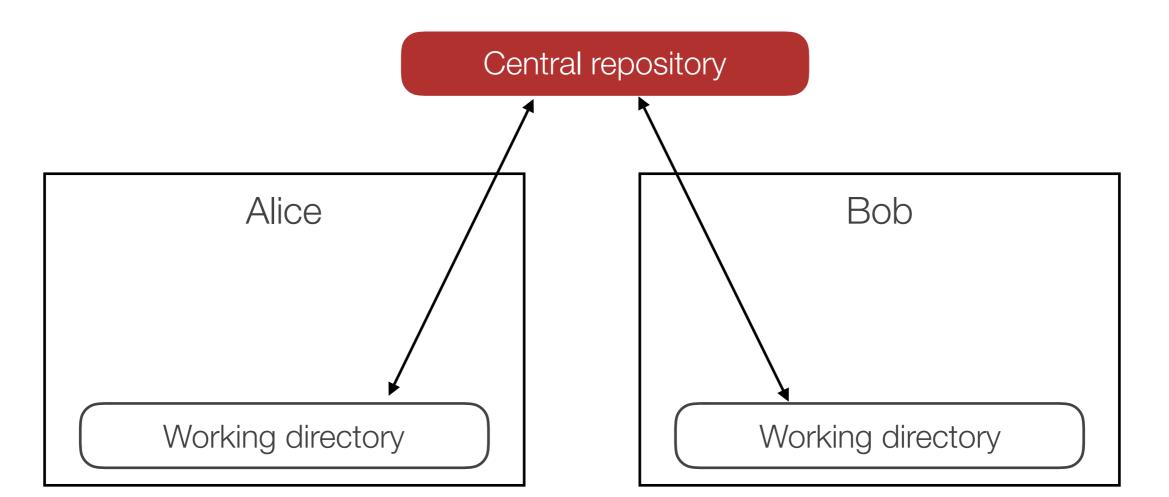


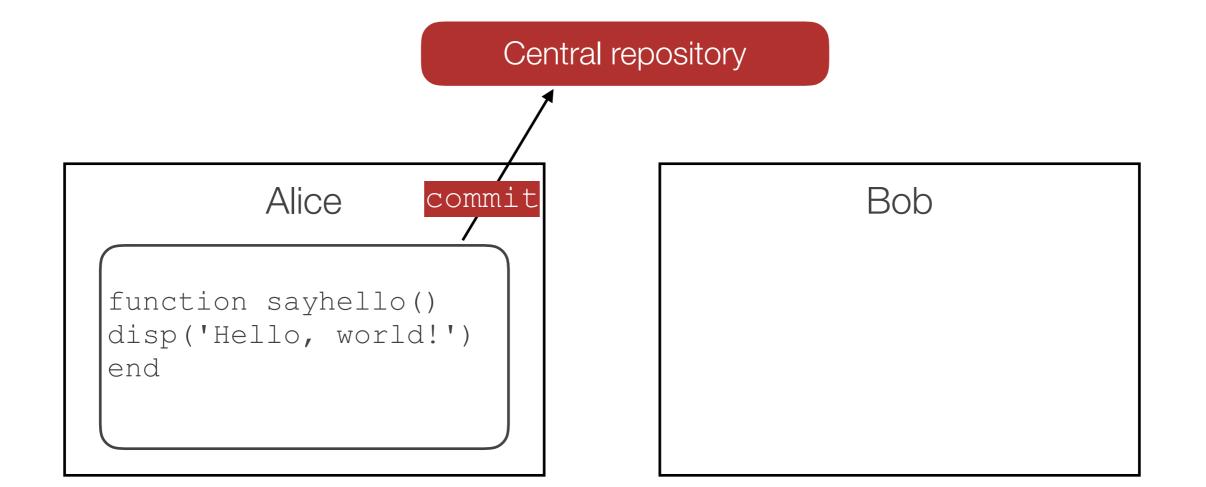
DONT #2: Never use a Centralized VCS

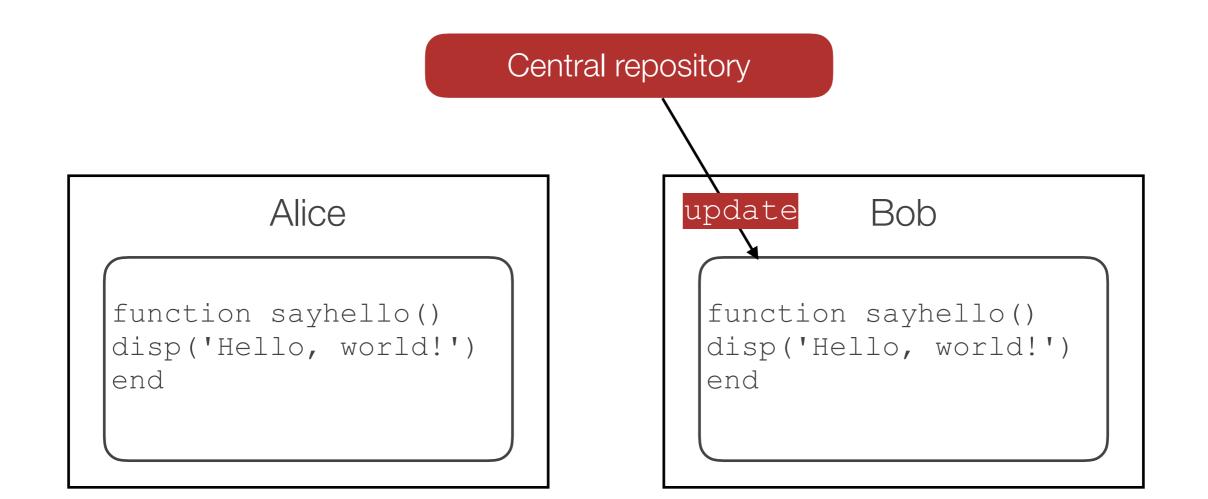
Main idea: use a centralized repository (public server)

Problems:

- requires permanent internet connection (+single point of failure)
- access rights management
- communication ("Hey, I made a change, you need to update")
- only one person can commit at a time

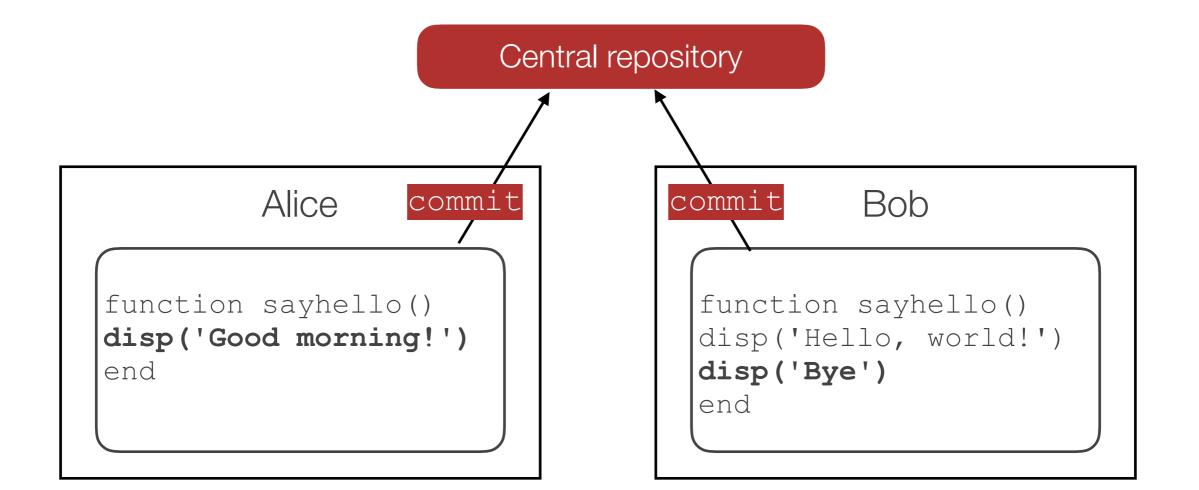






Centralized VCS will prevent simultaneous commits

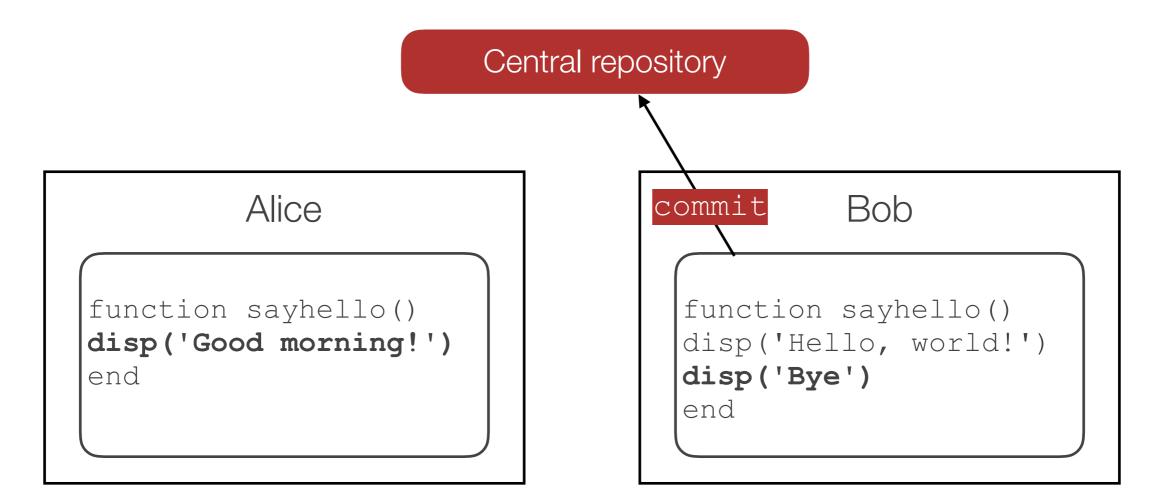
- in fact, "update" locks the repository for a particular user



Centralized VCS will prevent simultaneous commits

- in fact, "update" locks the repository for a particular user

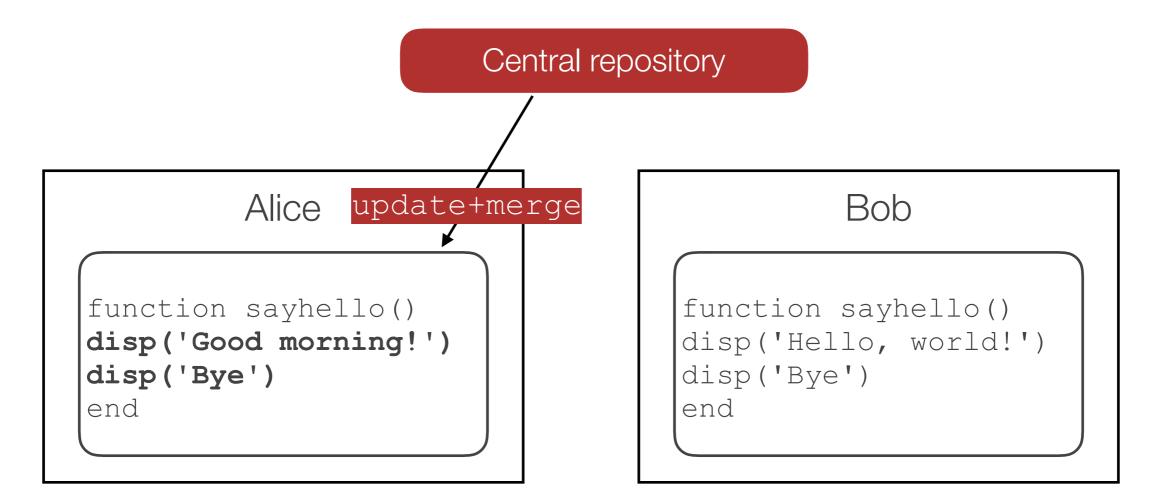
- Alice must wait for Bob to commit
- Alice then updates her working directory with Bob' changes
- Alice must "merge" Bob's changes with hers



Centralized VCS will prevent simultaneous commits

- in fact, "update" locks the repository for a particular user

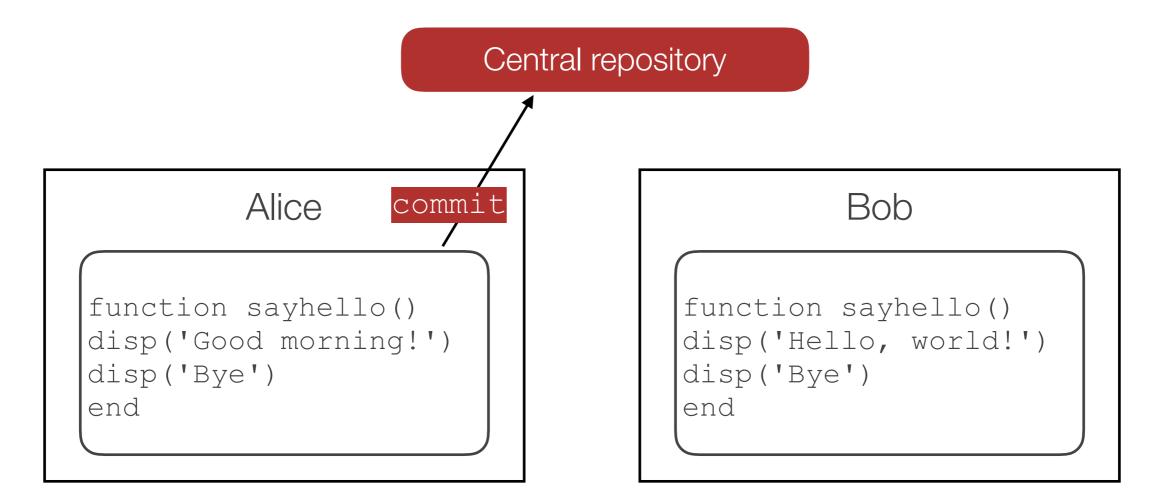
- Alice must wait for Bob to commit
- Alice then updates her working directory with Bob' changes
- Alice must "merge" Bob's changes with hers



Centralized VCS will prevent simultaneous commits

- in fact, "update" locks the repository for a particular user

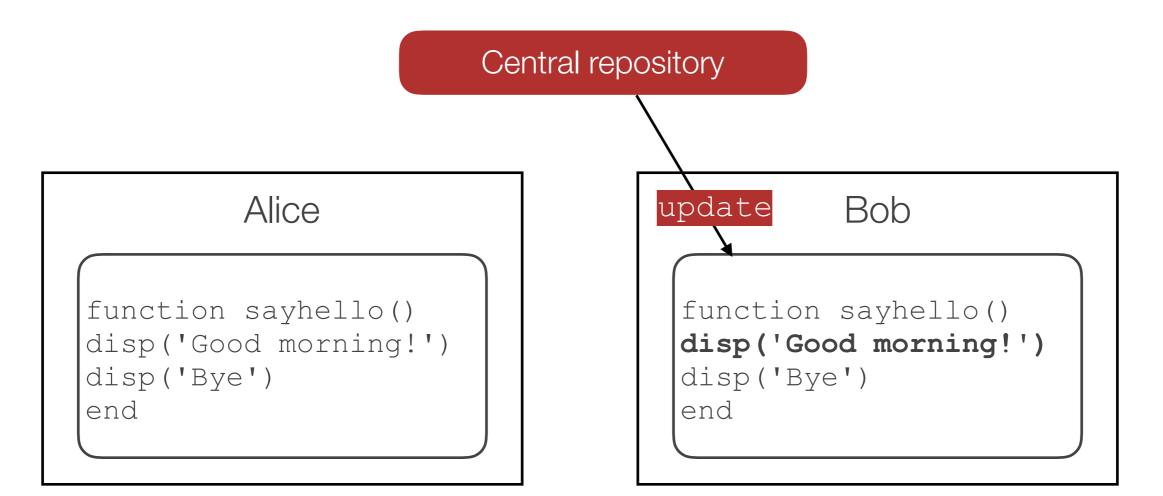
- Alice must wait for Bob to commit
- Alice then updates her working directory with Bob' changes
- Alice must "merge" Bob's changes with hers



Centralized VCS will prevent simultaneous commits

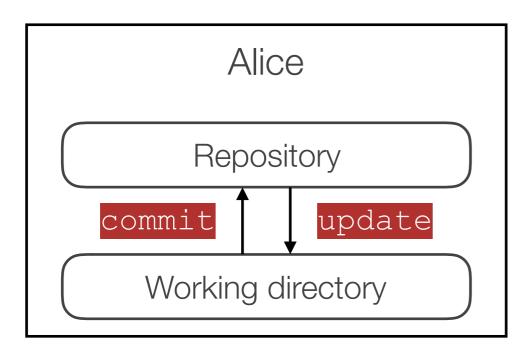
- in fact, "update" locks the repository for a particular user

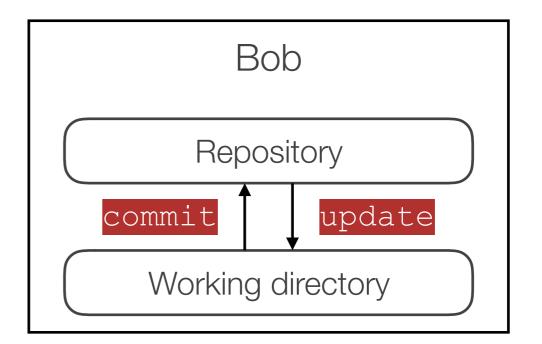
- Alice must wait for Bob to commit
- Alice then updates her working directory with Bob' changes
- Alice must "merge" Bob's changes with hers



Distributed VCS (Mercurial, Git)

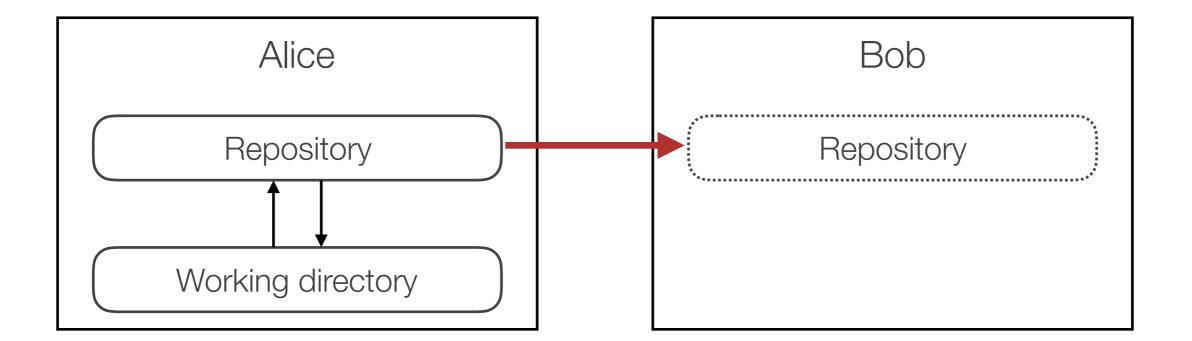
Main idea: everybody has a local copy of the repository

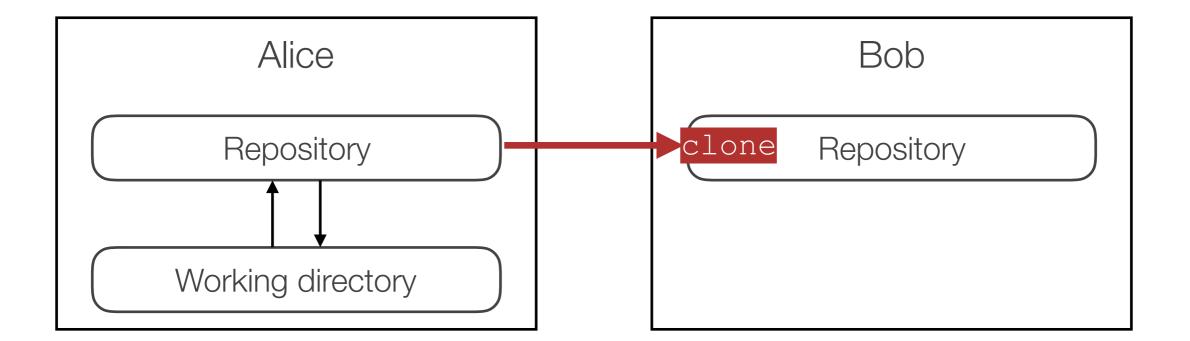


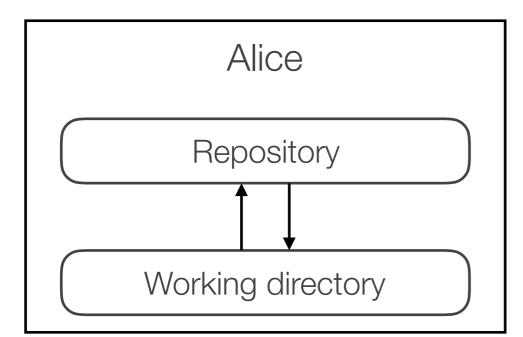


Distributed VCS (Mercurial, Git)

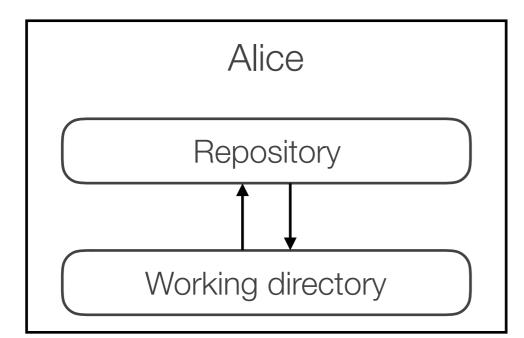
Main idea: everybody has a local copy of the repository

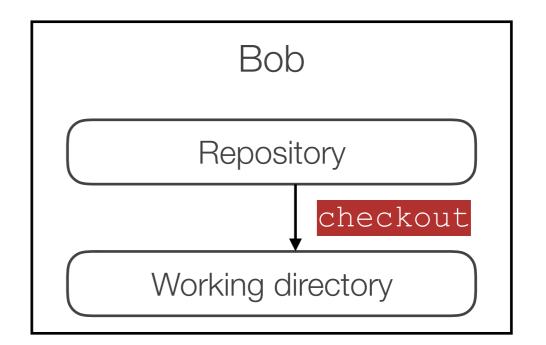




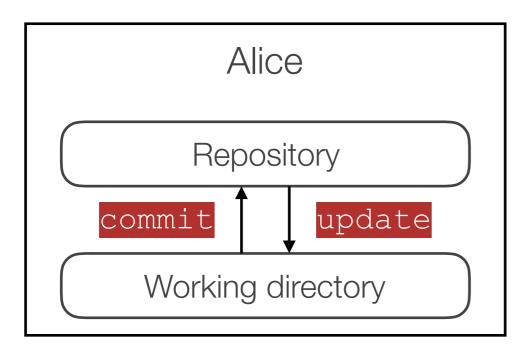


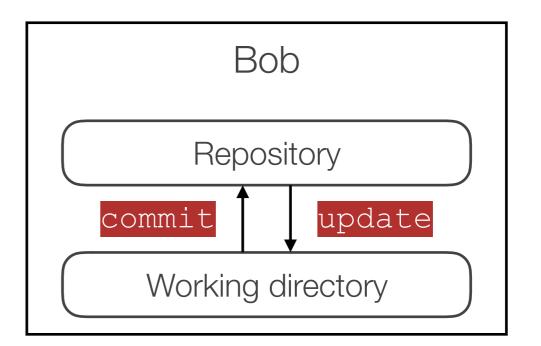
	Bob	
	Repository	
**************************************	Working directory	***



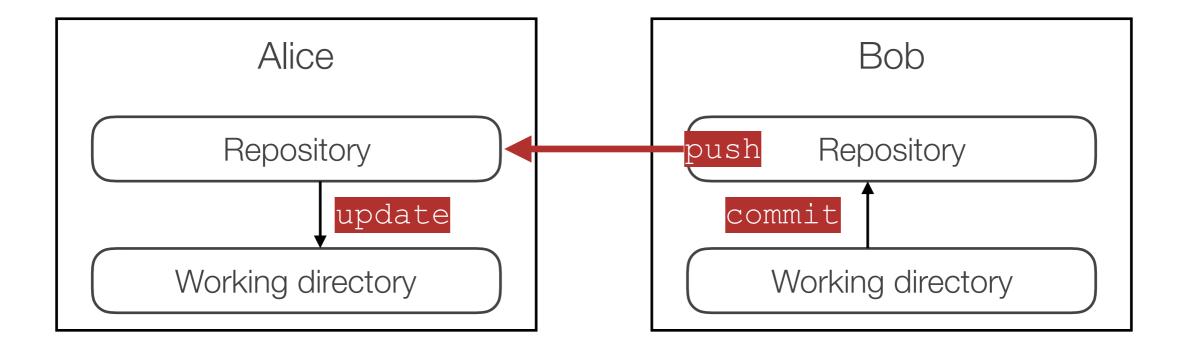


- every developer can commit/update at any time to his/her repository
- no permanent connection required

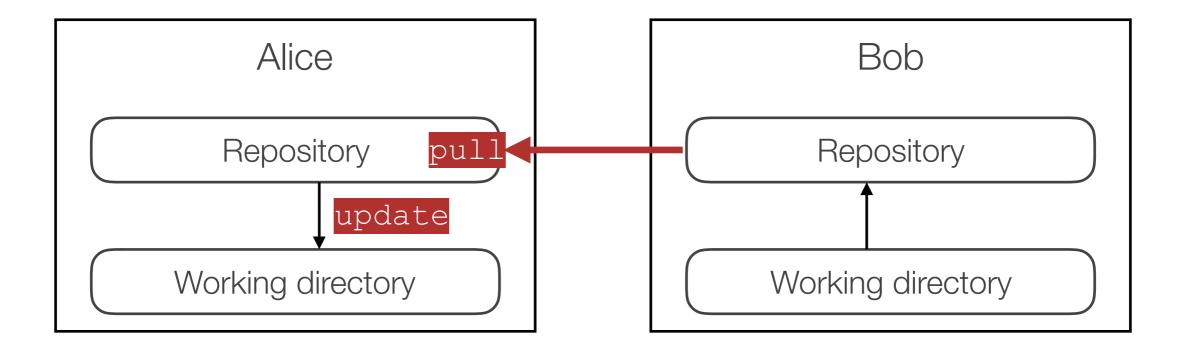




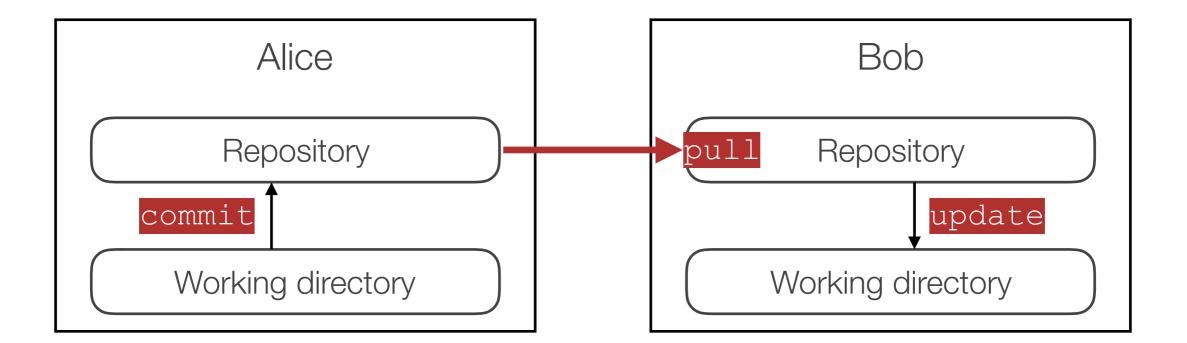
- every developer can commit/update at any time to his/her repository
- no permanent connection required



- every developer can commit/update at any time to his/her repository
- no permanent connection required



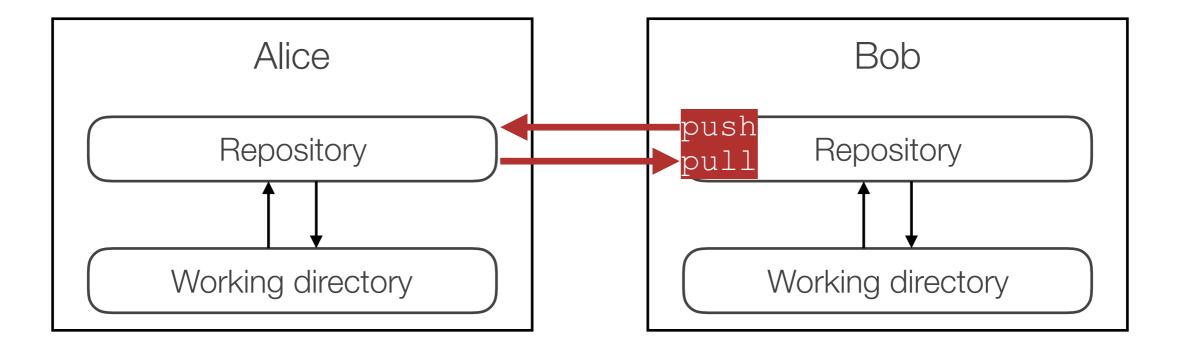
- every developer can commit/update at any time to his/her repository
- no permanent connection required



Main idea: everybody has a local copy of the repository

- every developer can commit/update at any time to his/her repository
- no permanent connection required

Rule of thumb: never push to a remote repository, let its owner pull from yours

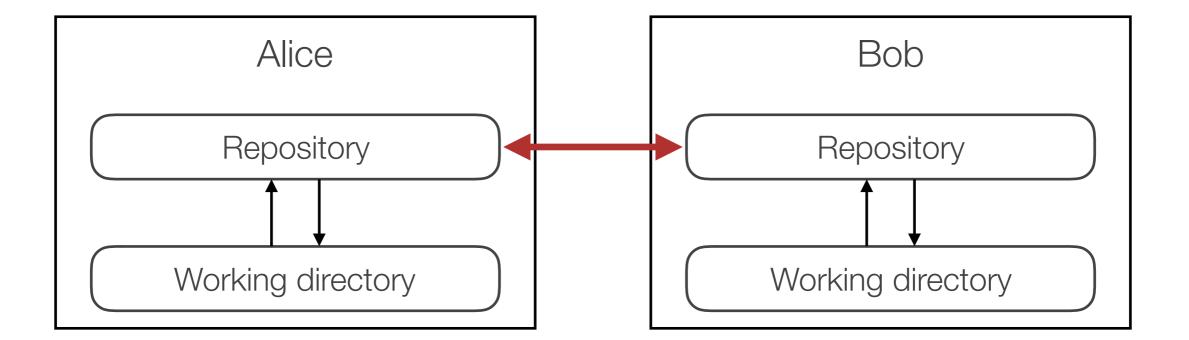


Main idea: everybody has a local copy of the repository

- every developer can commit/update at any time to his/her repository
- no permanent connection required

Questions:

- access rights management?
- communication ("Hey, I made a change, you need to update")?
- which repository is the "blessed" one for production code?

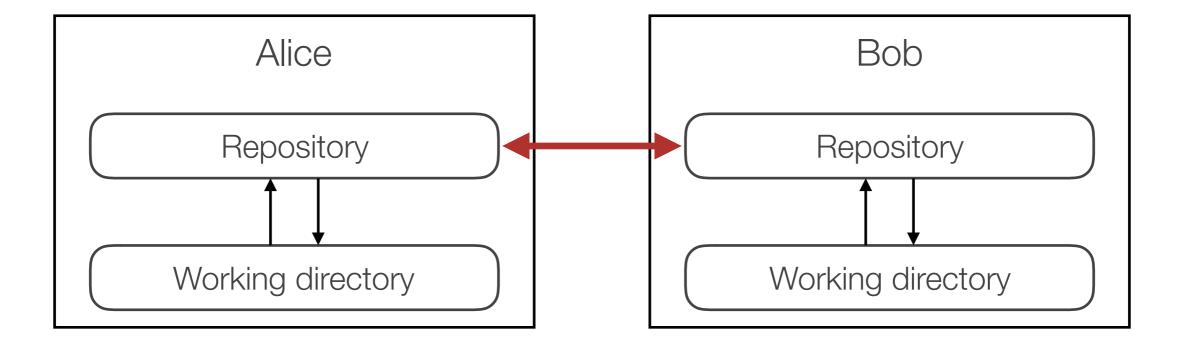


Main idea: everybody has a local copy of the repository

- every developer can commit/update at any time to his/her repository
- no permanent connection required

Questions:

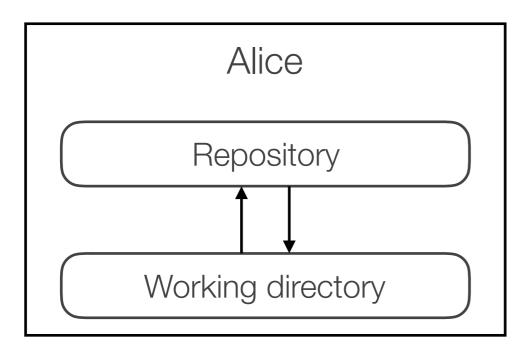
- access rights management?
- communication ("Hey, I made a change, you need to update")?
- which repository is the "blessed" one for production code?

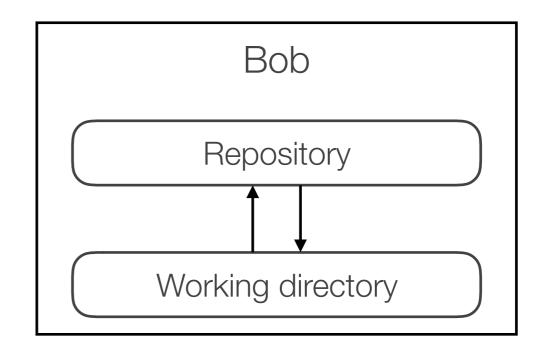


Distributed VCS Workflows

Three possible workflows:

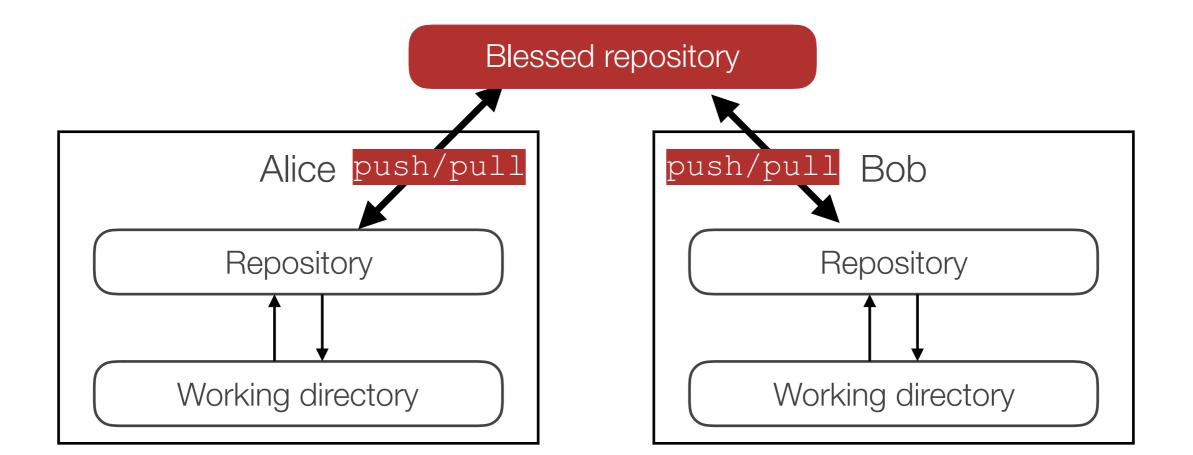
- centralized
- integration manager
- dictator / lieutenants





Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)



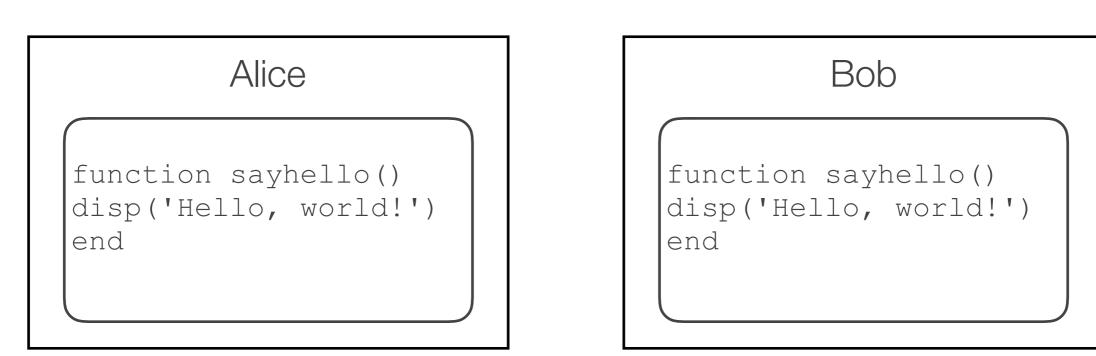
Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

Problems:

- cannot push conflicting changes

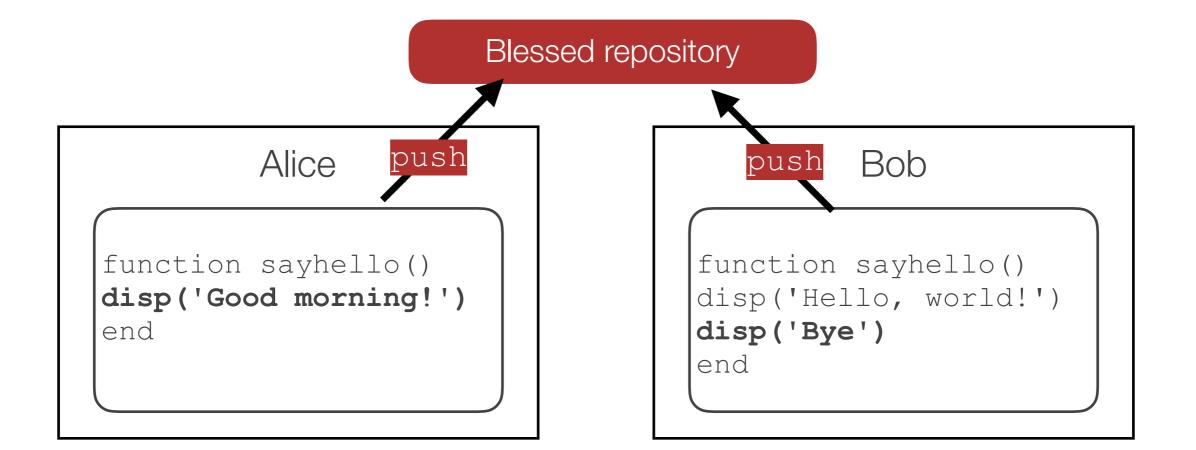
Blessed repository



Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

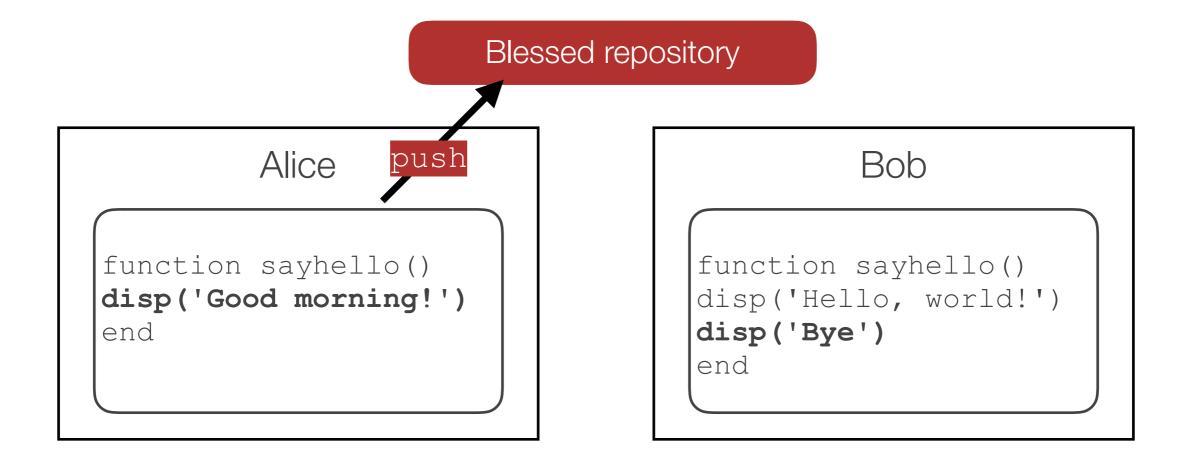
Problems:



Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

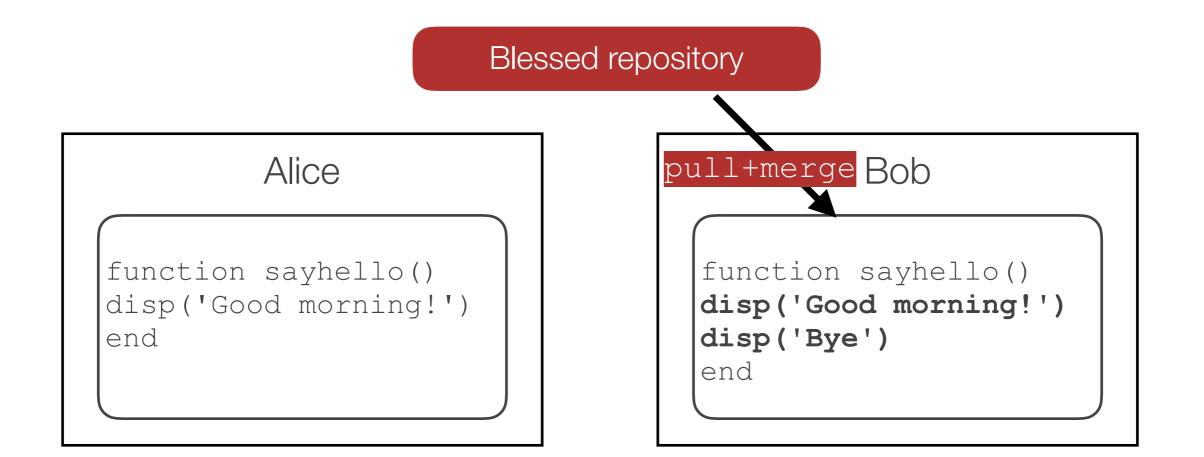
Problems:



Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

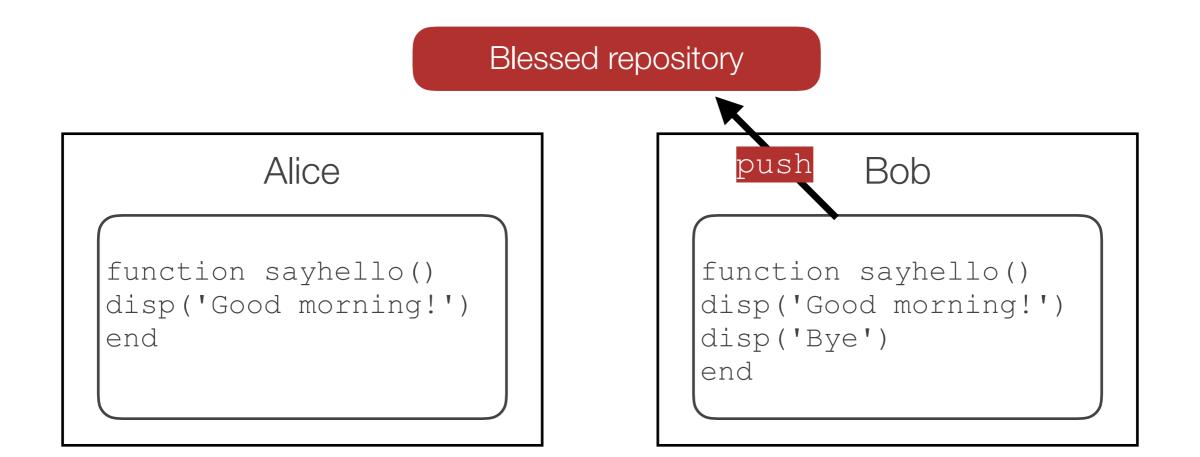
Problems:



Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

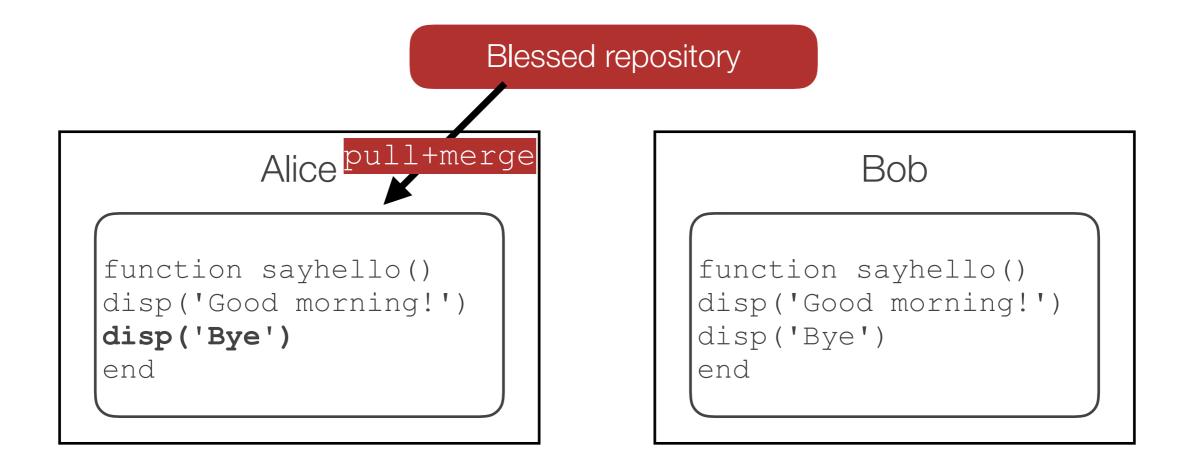
Problems:



Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

Problems:

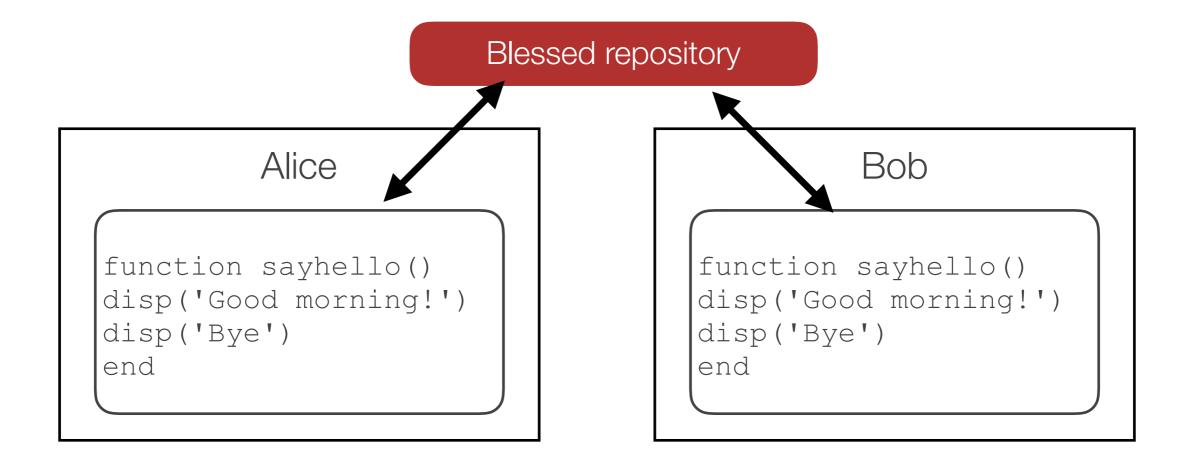


Idea: use a central "blessed" repository

- but each developer has his/her own copy (no single point of failure)

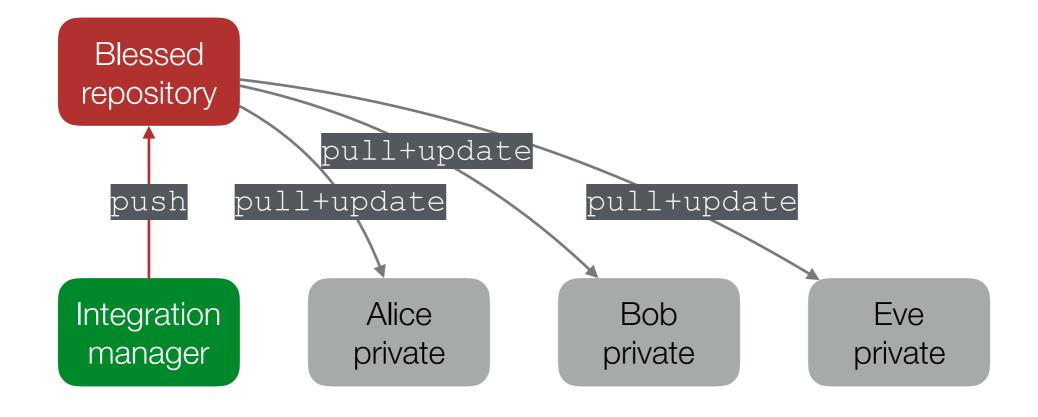
Problems:

- cannot push conflicting changes (merging is required, often difficult!)
- only suitable for a small number of developers (<5)



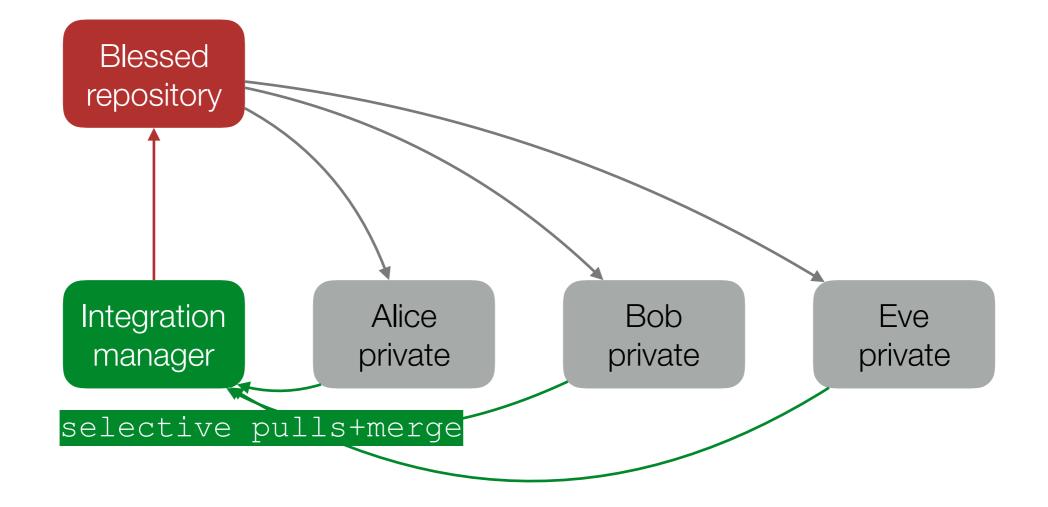
Idea: use a central "blessed" repository:

- but only the integration manager can push to it



Idea: use a central "blessed" repository:

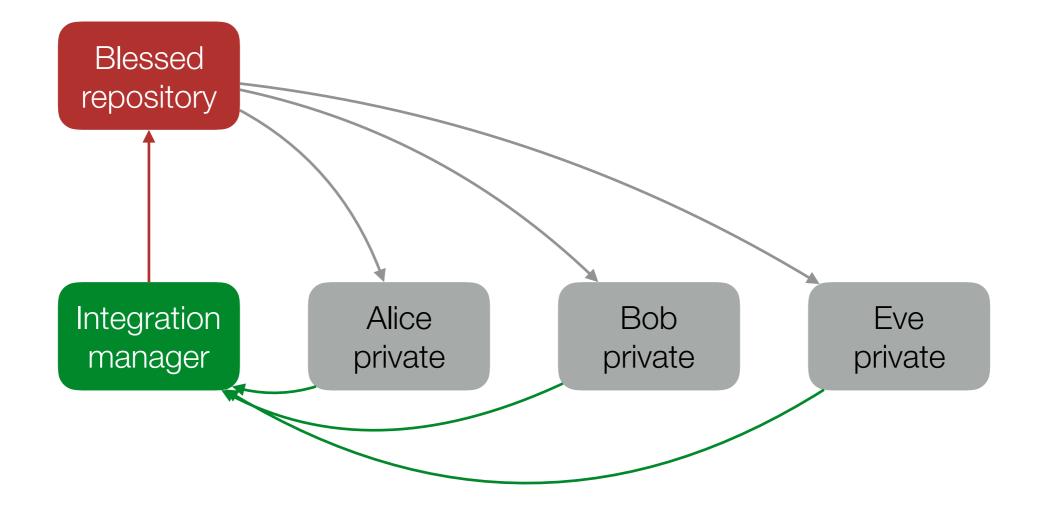
- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges



Idea: use a central "blessed" repository:

- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges

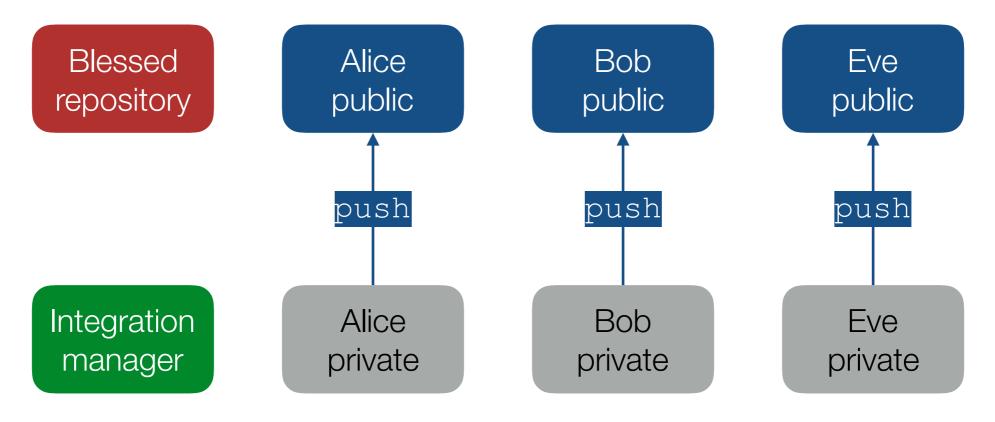
Problem: usually cannot pull from private repositories



Idea: use a central "blessed" repository:

- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges

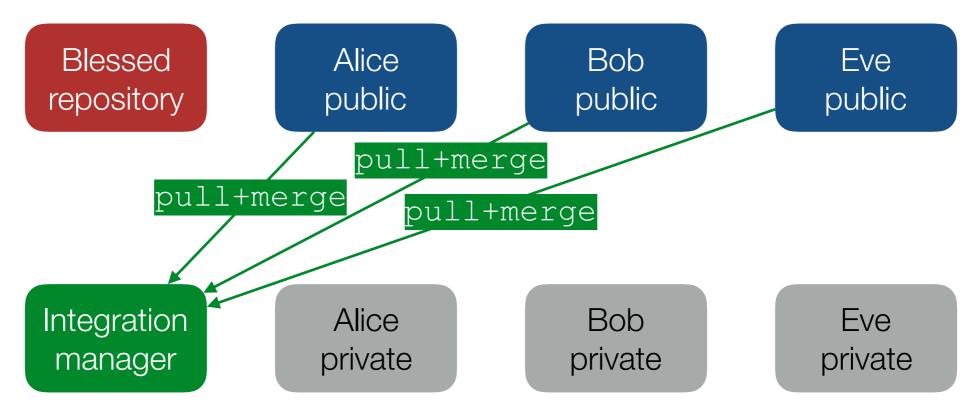
Problem: usually cannot pull from private repositories



Idea: use a central "blessed" repository:

- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges

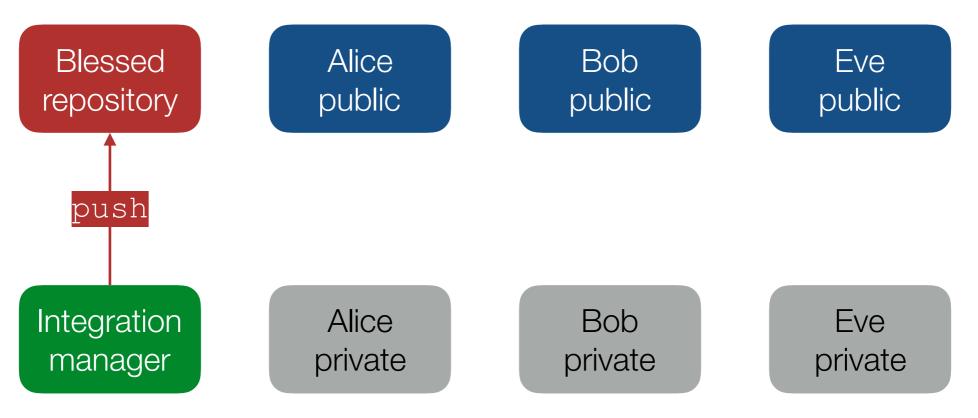
Problem: usually cannot pull from private repositories



Idea: use a central "blessed" repository:

- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges

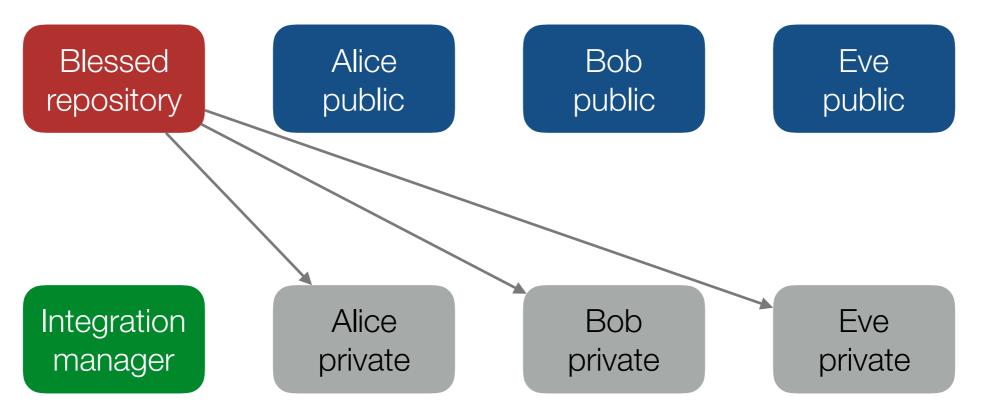
Problem: usually cannot pull from private repositories



Idea: use a central "blessed" repository:

- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges

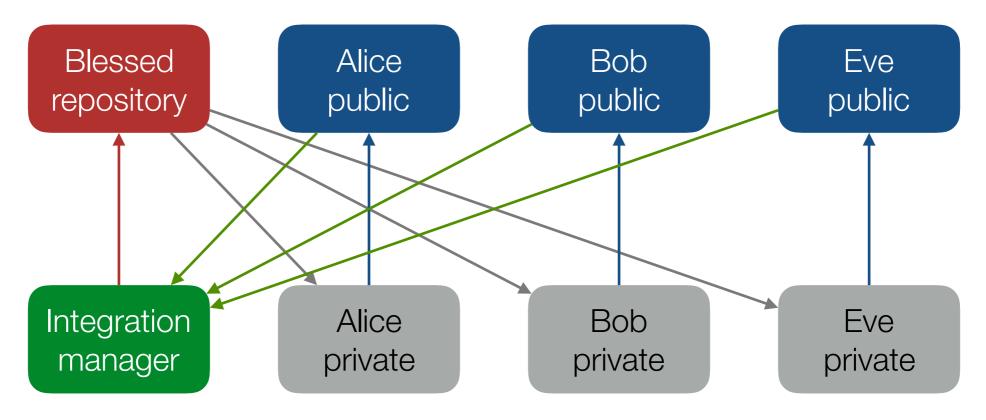
Problem: usually cannot pull from private repositories



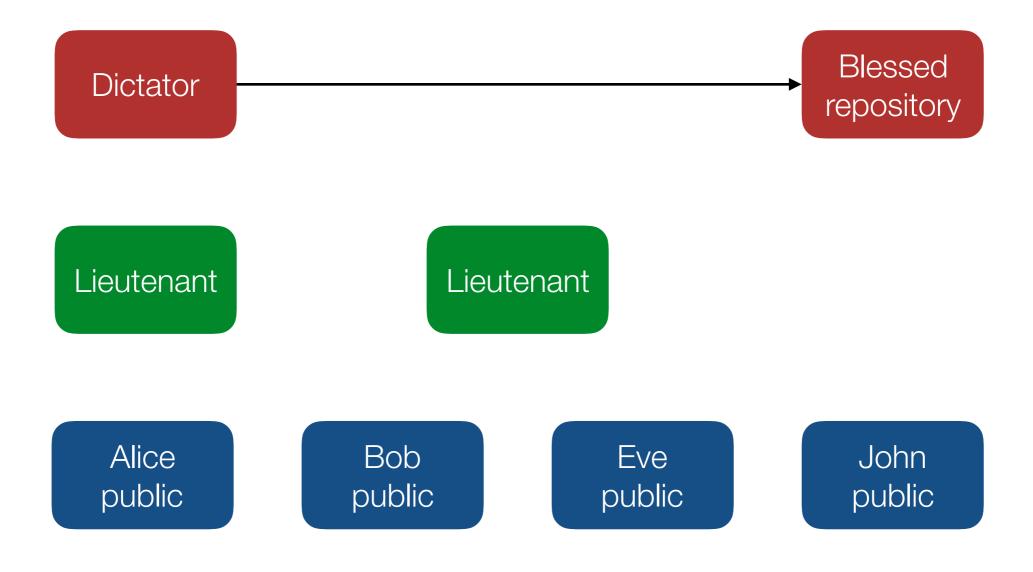
Idea: use a central "blessed" repository:

- but only the integration manager can push to it
- the manager cherry-picks changes and handles merges

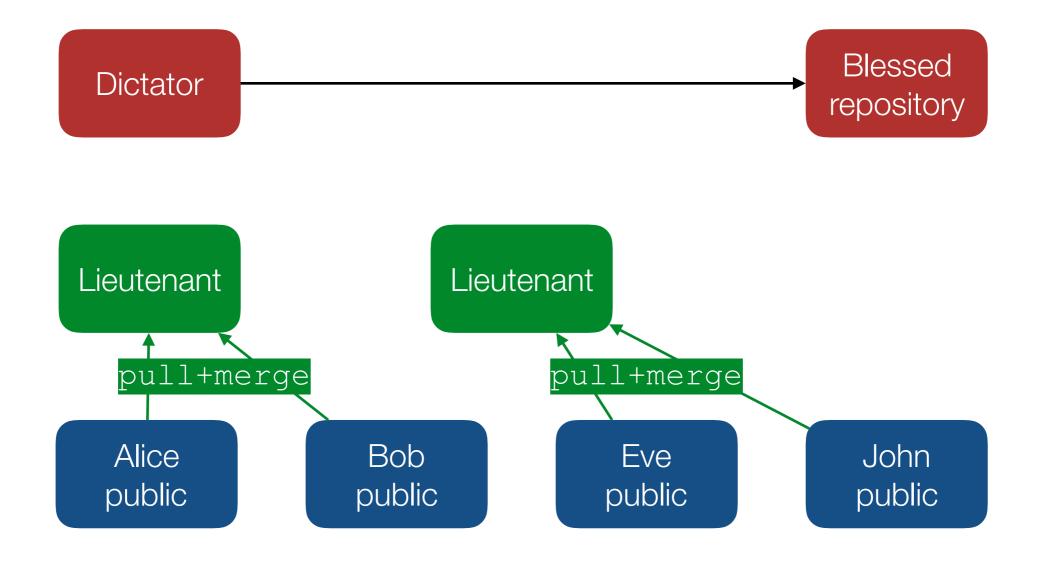
Problem: usually cannot pull from private repositories



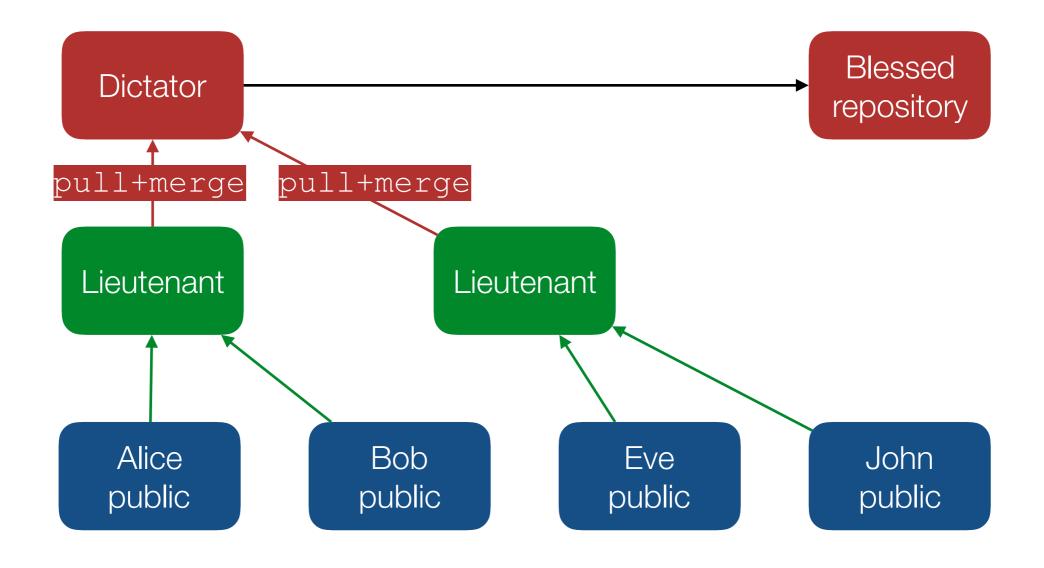
- each lieutenant is usually responsible for one module
- used for big projects (e.g. the Linux kernel, Python)



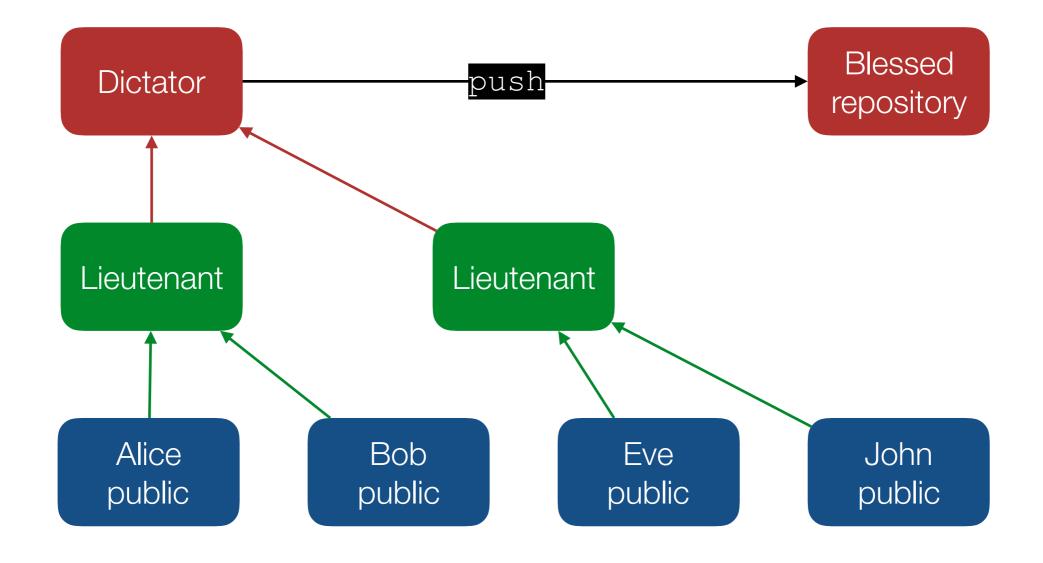
- each lieutenant is usually responsible for one module
- used for big projects (e.g. the Linux kernel, Python)



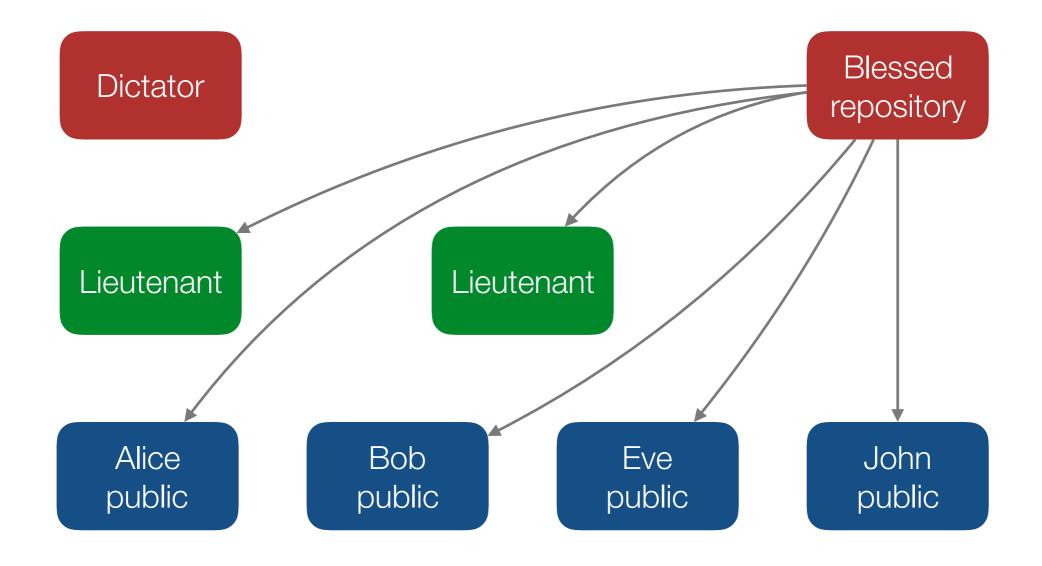
- each lieutenant is usually responsible for one module
- used for big projects (e.g. the Linux kernel, Python)



- each lieutenant is usually responsible for one module
- used for big projects (e.g. the Linux kernel, Python)



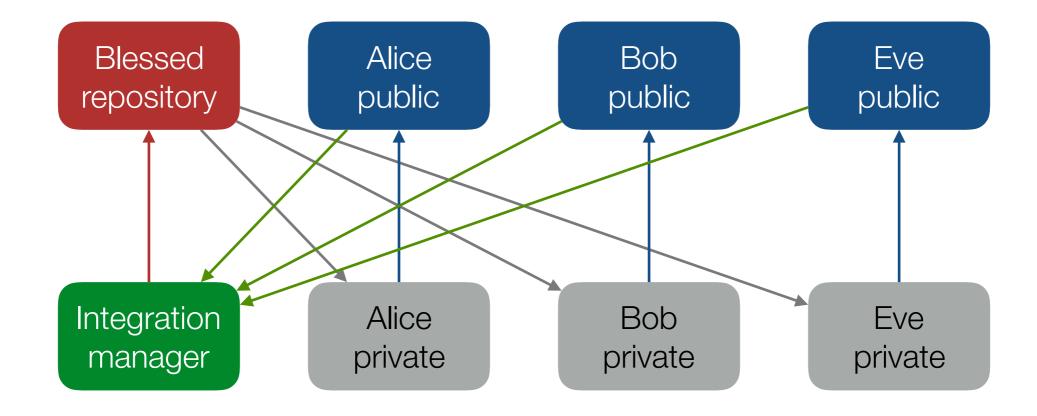
- each lieutenant is usually responsible for one module
- used for big projects (e.g. the Linux kernel, Python)



DO #2: Use the Integration Manager Workflow

Main objectives:

- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development



DO #3: Choose a Proper Development Model

Main objectives:

- record changes over time
- recall a specific version later
- enable collaboration
- allow nonlinear development

DO #3: Choose a Proper Development Model

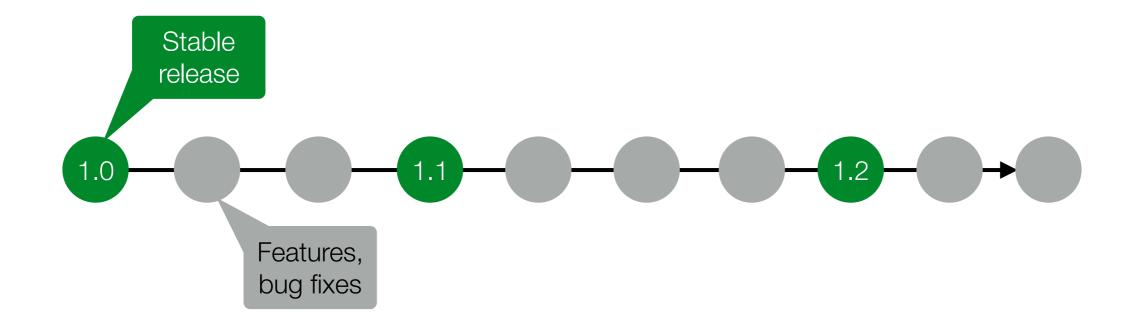
Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

DO #3: Choose a Proper Development Model

Development models:

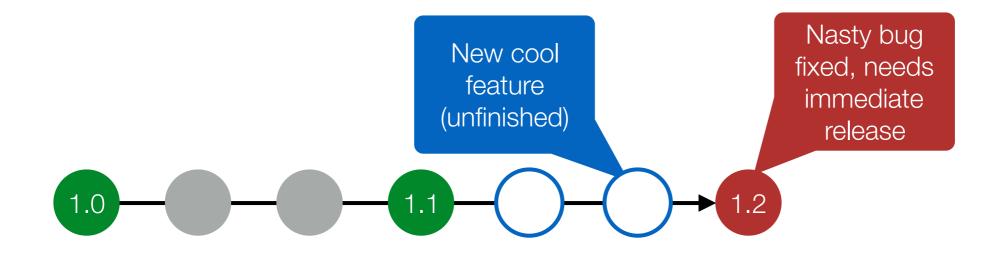
- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model



Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

Problem: release with unfinished features is not desired

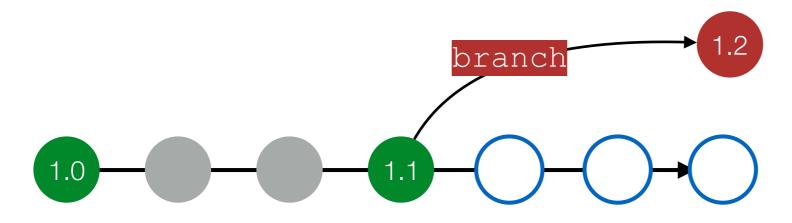


Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

Problem: release with unfinished features is not desired

Solution: branch from latest stable release, only merge with finished features

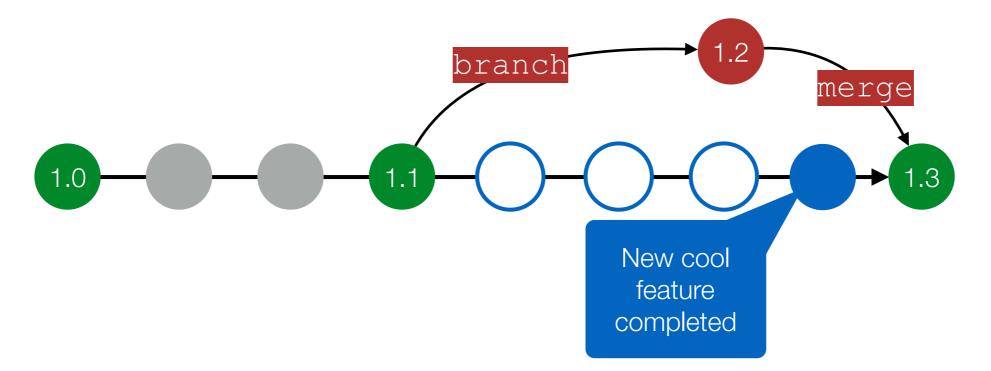


Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

Problem: release with unfinished features is not desired

Solution: branch from latest stable release, only merge with finished features

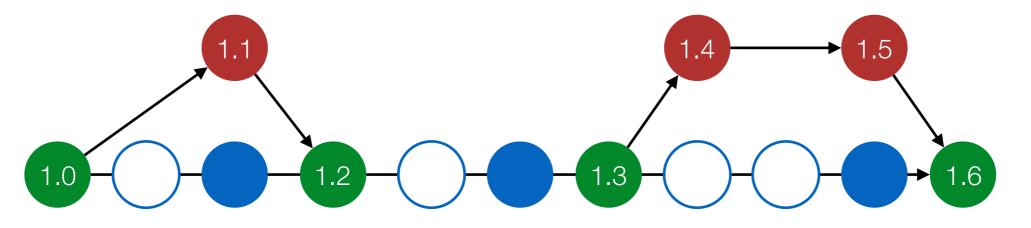


Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

Problem: release with unfinished features is not desired

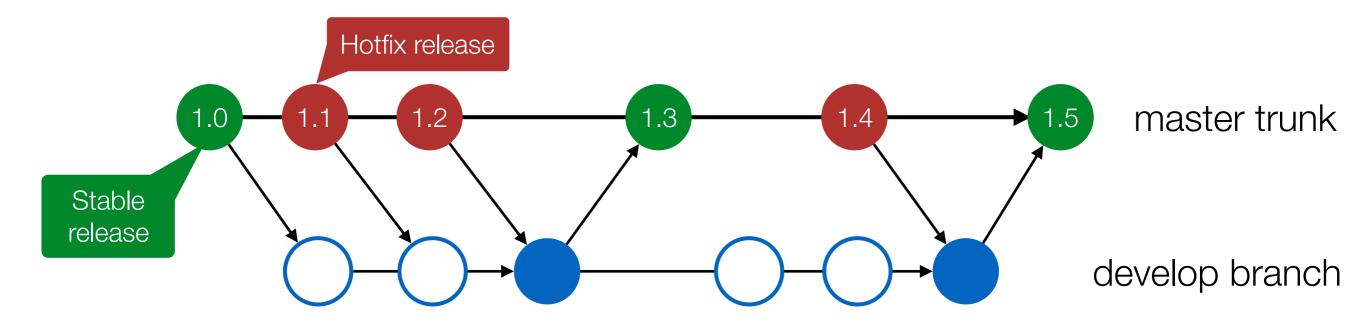
Solution: branch from latest stable release, only merge with finished features (but then the production code is the second-class citizen)



Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

Idea: the trunk always contains production-ready code

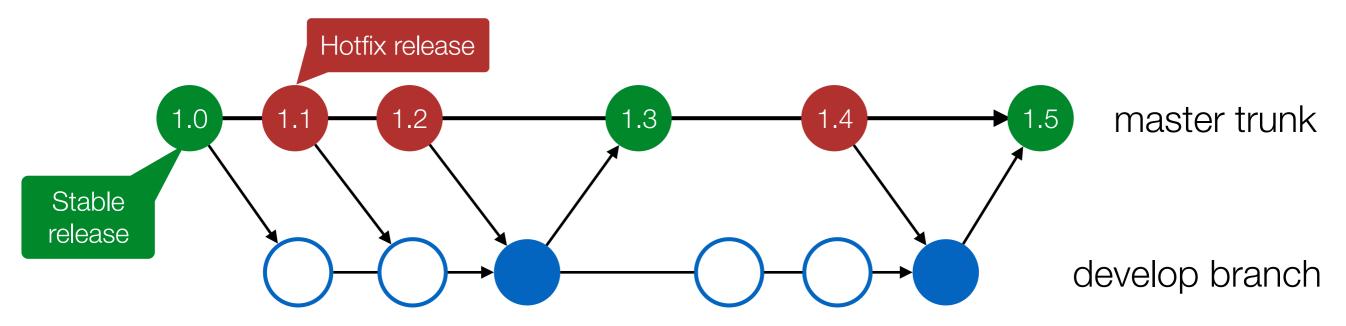


Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model

Idea: the trunk always contains production-ready code

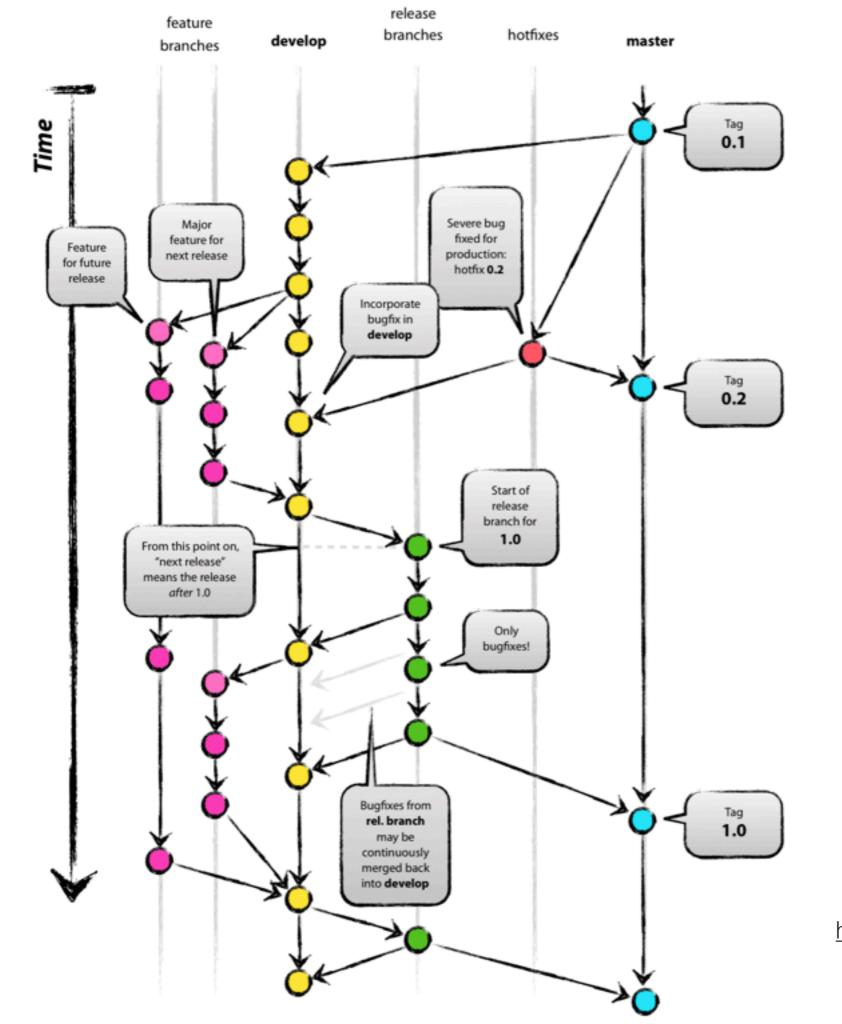
Problem: what if we work on multiple new features in parallel?



Development models:

- linear
- hotfix branches
- "master" trunk and "develop" branch
- Driessen's branching model





http://nvie.com/posts/ <u>a-successful-git-</u> <u>branching-model/</u>

Development models:

- linear
- hotfix branches for projects with infrequent changes
- "master" trunk and "develop" branch for a single developer
- Driessen's branching model for projects with many contributors

Don't be afraid of branches, modern VCS's allow to "close and hide" branches (even unfinished ones)

Many open-source projects will only pull from you if your change is in a feature branch

DOs and DONTs

DONTs:

- do not use archives as a substitute for version control
- never use a centralized VCS (CVS, SVN)
- do not store automatically generated files (PDF, binaries, etc.) and sensitive data (e.g. passwords, logins, SSH keys, etc.)

DOs:

- use a distributed VCS (Mercurial, Git)
- use the "Integration Manager" workflow (although the "centralized" workflow works well for papers)
- use Driessen's branching model

Agenda

Today

- version control systems and collaborative development
- Mercurial, Git, Bitbucket, GitHub
- providing support

Tomorrow

- unit testing
- documentation
- dissemination

Mercurial vs Git



Both written in April 2005 after the big BitKeeper/Torvalds split

Both are:

- distributed version control systems
- fast (Git a tad faster)
- relatively easy to use once (Mercurial more so)
- multi-platform (Mercurial "more native" on Windows)

Selecting one is mostly a matter of taste

- Git is more popular mainly due to GitHub

Mercurial vs Git



Bitbucket	Octave
Dovecot	OpenJDK
FuseSMB	OpenOffice
Growl	OpenSolaris
MPT	Python
mutt	RabbitMQ
NetBeans	rpm.org
nginx	Sphinx



Android Debian tools Drupal FFmpeg GCC GNOME jQuery Julia KVM Linux kernel Maemo

OLPC

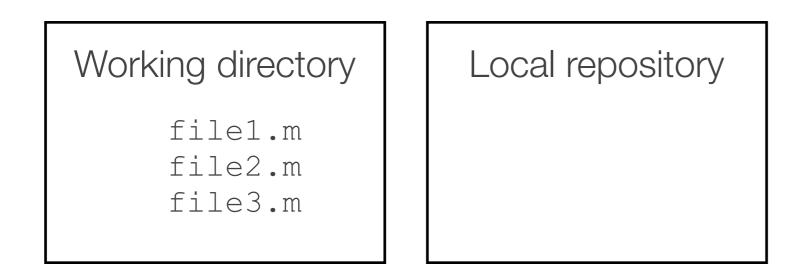
Perl

Samba

VLC

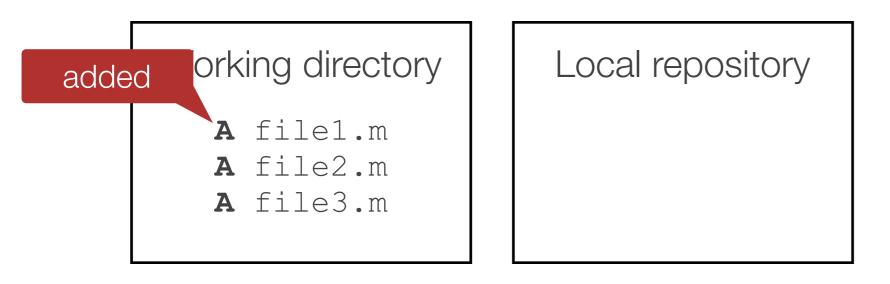
Yum

```
Working directory
file1.m
file2.m
file3.m
```

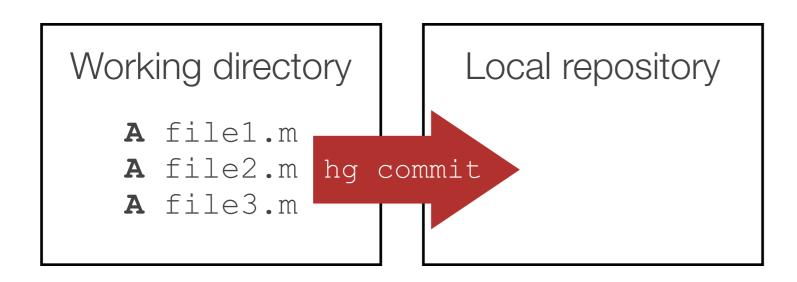




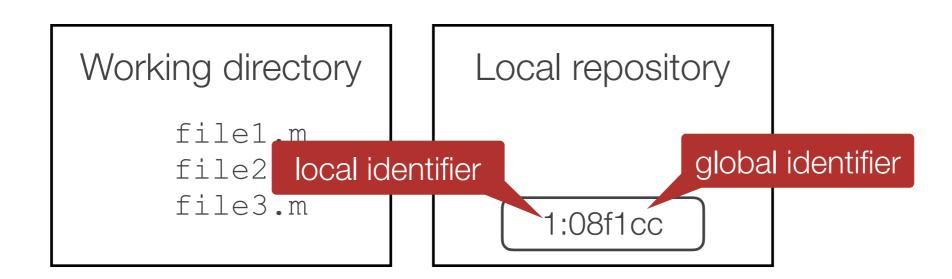
Initialize a new local repository: hg init Check status of files: hg status



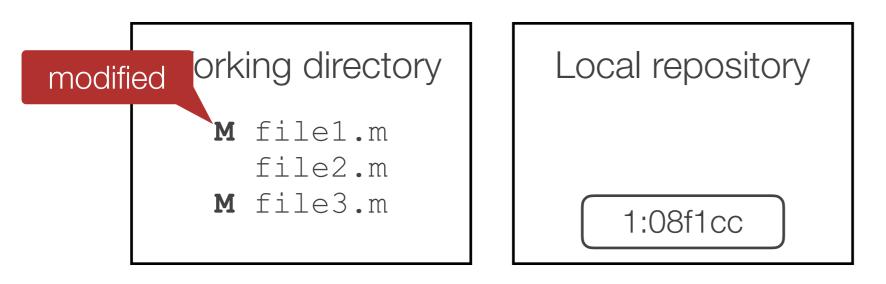
Initialize a new local repository: hg init Check status of files: hg status Start tracking all files: hg add



- Check status of files: hg status
- Start tracking all files: hg add
- Commit the added files to the repository: hg commit -m "message"



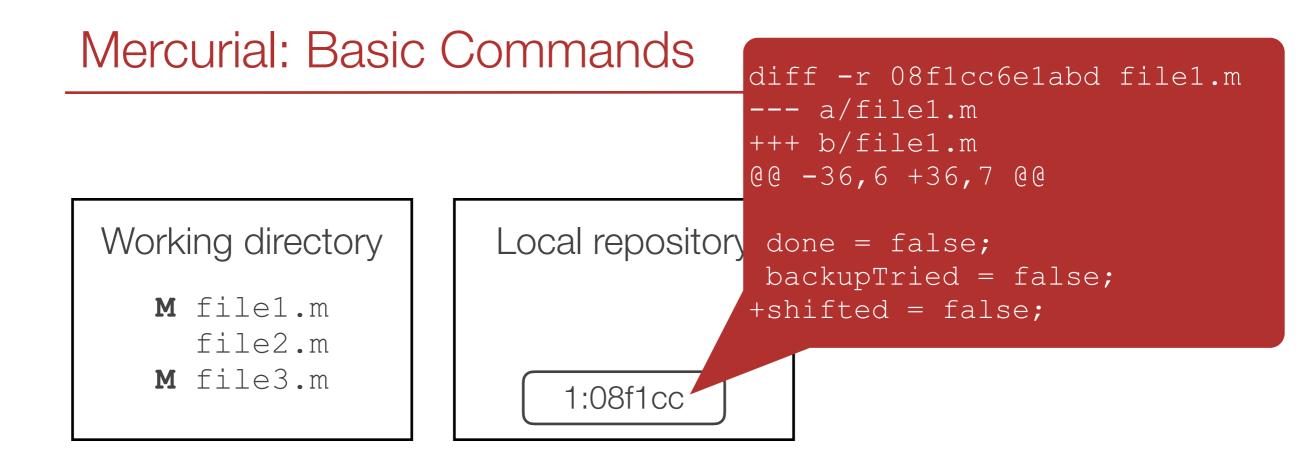
- Check status of files: hg status
- Start tracking all files: hg add
- Commit the added files to the repository: hg commit -m "message"



Check status of files: hg status

Start tracking all files: hg add

Commit the added files to the repository: hg commit -m "message"

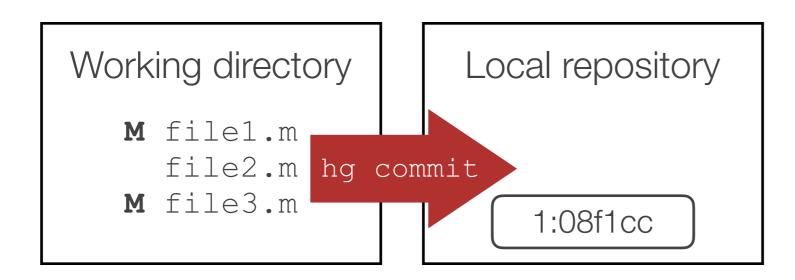


Check status of files: hg status

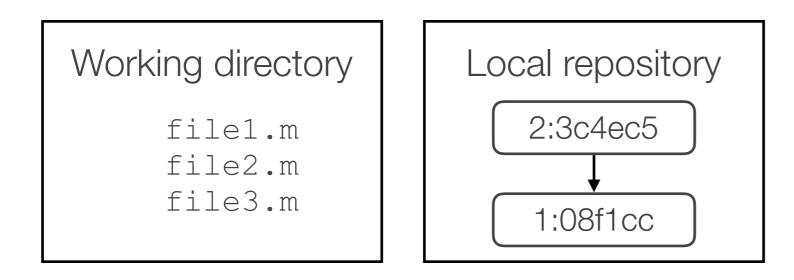
Start tracking all files: hg add

Commit the added files to the repository: hg commit -m "message"

View differences to repository: hg diff



- Check status of files: hg status
- Start tracking all files: hg add
- Commit the added files to the repository: hg commit -m "message"
- View differences to repository: hg diff
- Commit the modified files to the repository: hg commit -m "message"



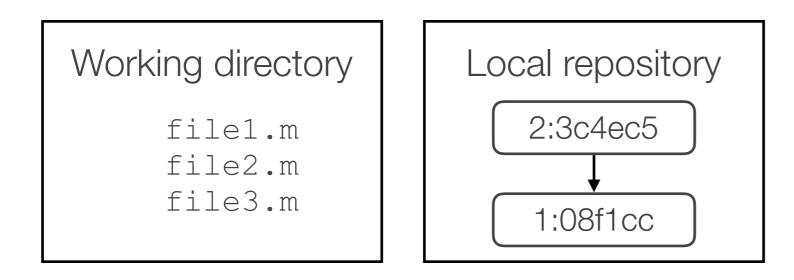
Check status of files: hg status

Start tracking all files: hg add

Commit the added files to the repository: hg commit -m "message"

View differences to repository: hg diff

Commit the modified files to the repository: hg commit -m "message"



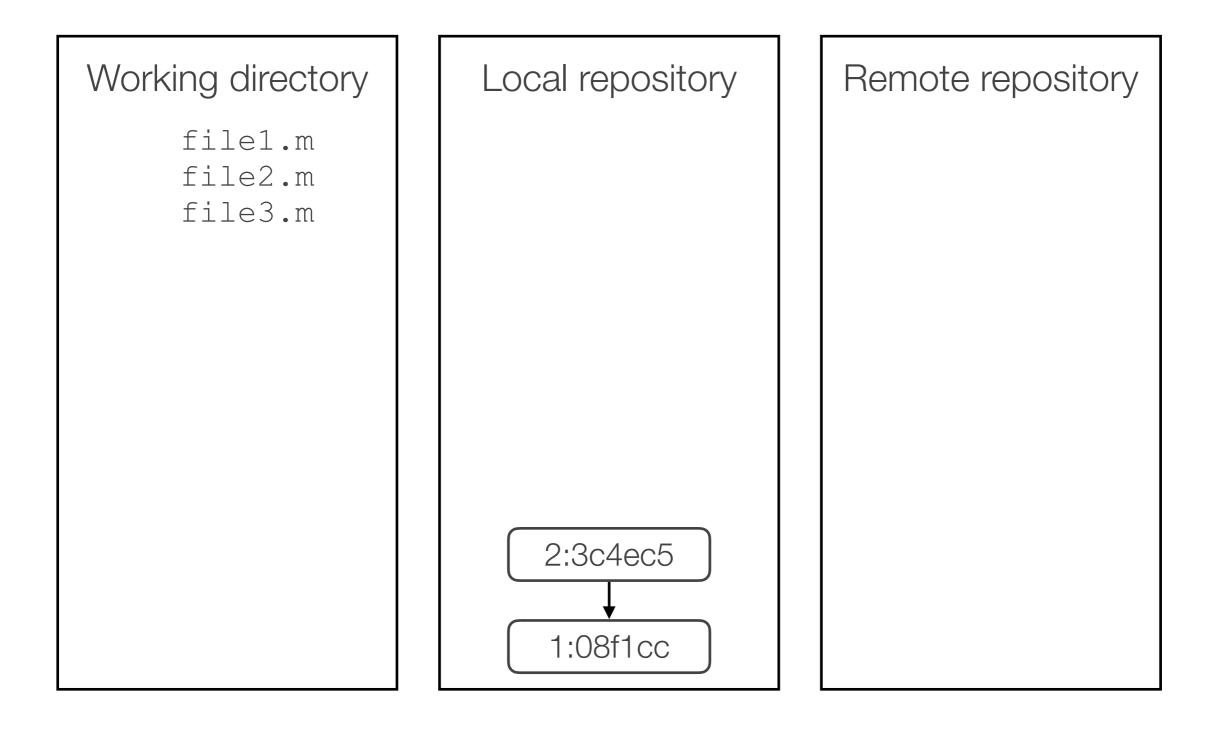
Check status of files: hg status

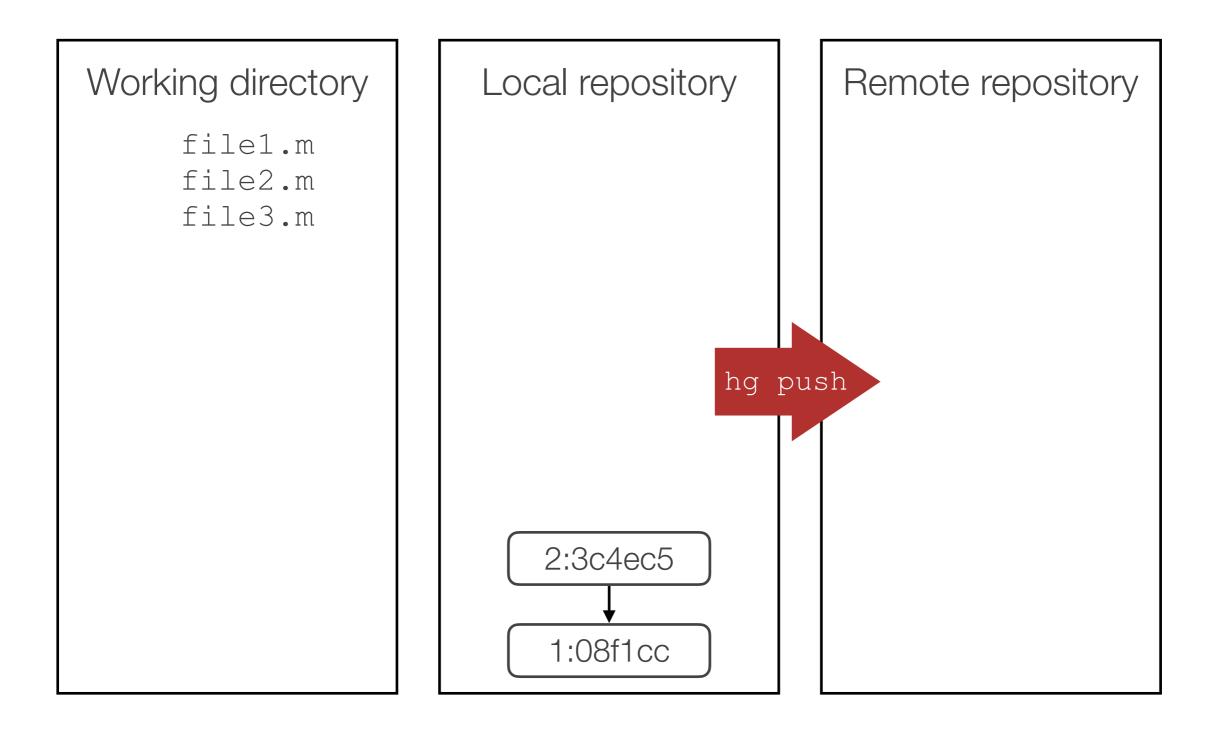
Start tracking all files: hg add

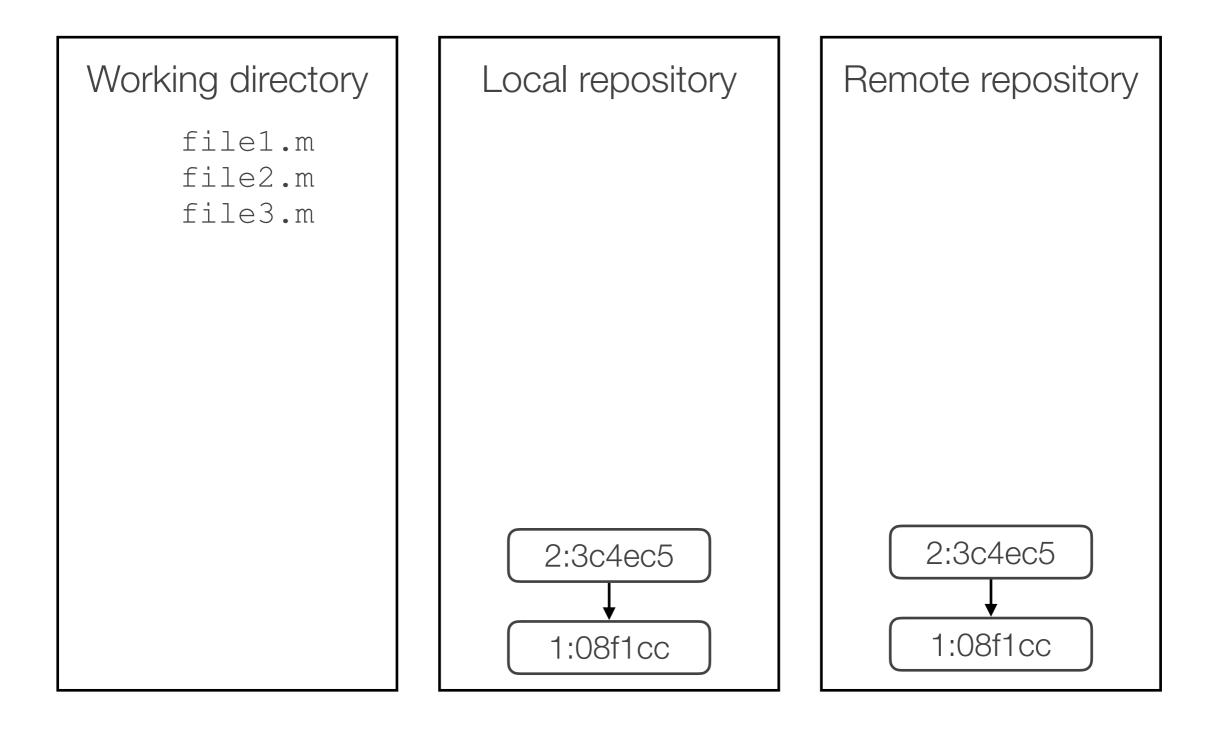
Commit the added files to the repository: hg commit -m "message"

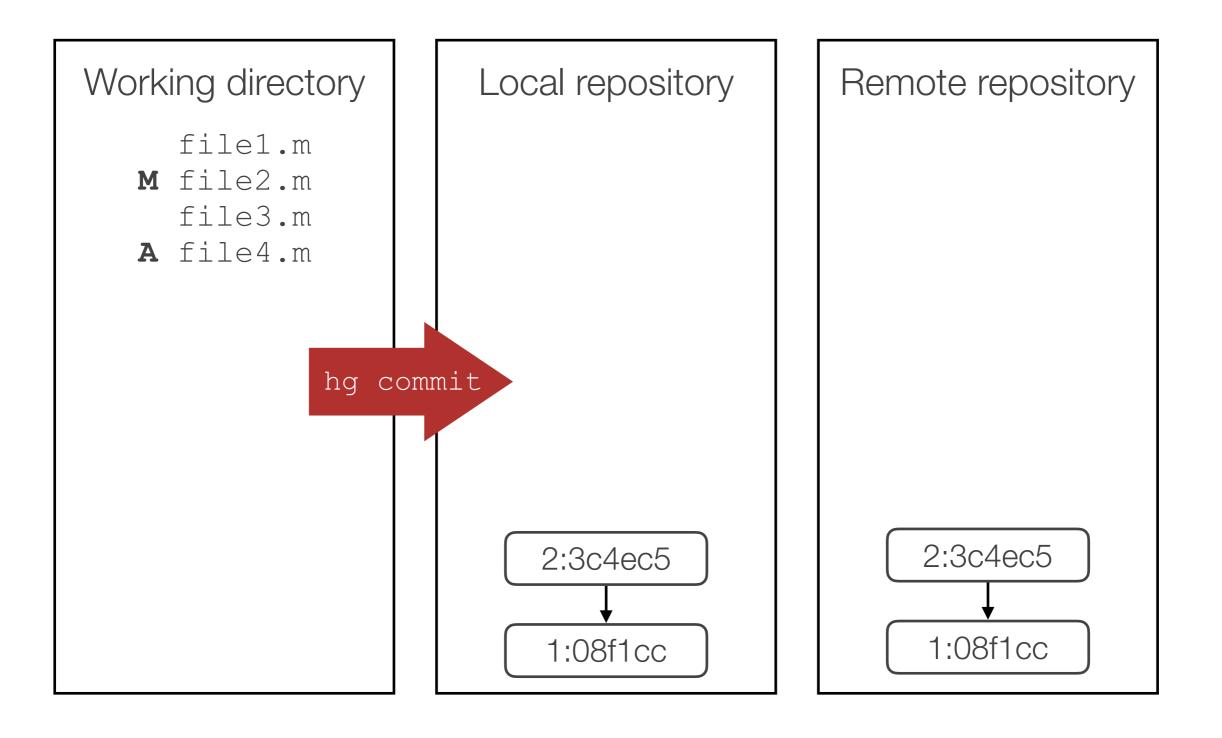
View differences to repository: hg diff

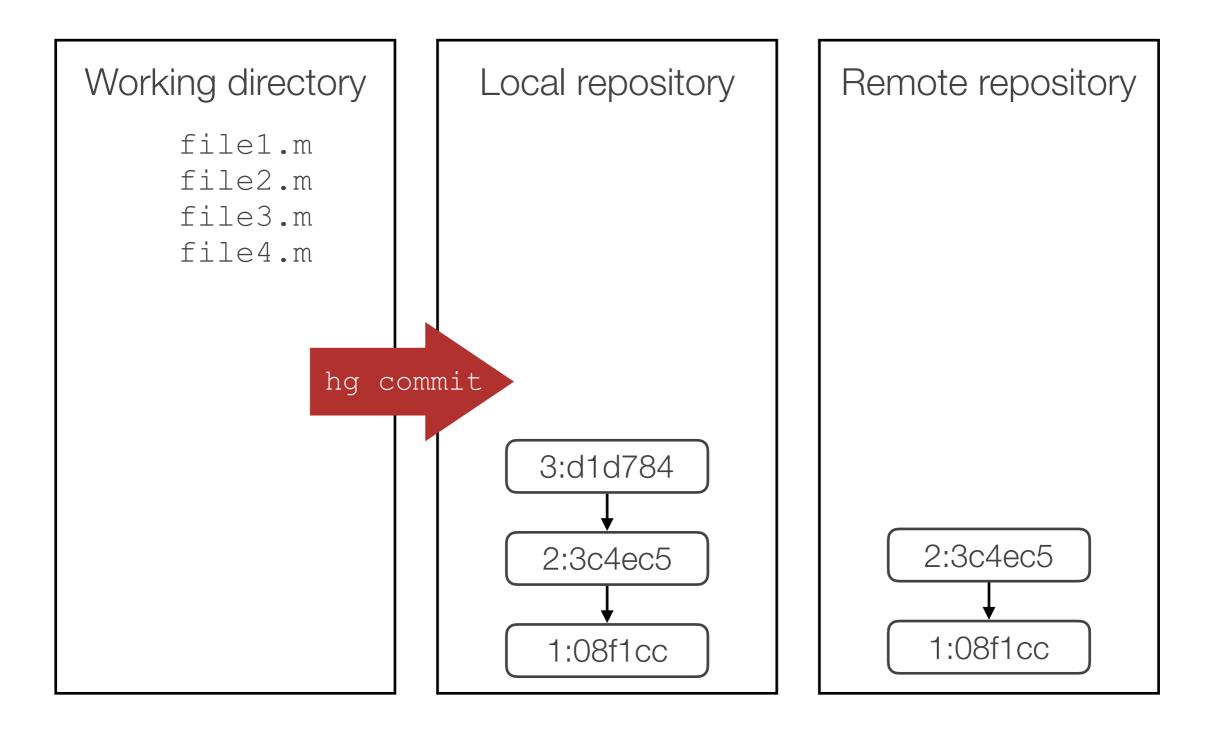
Commit the modified files to the repository: hg commit -m "message" View the log: hg log

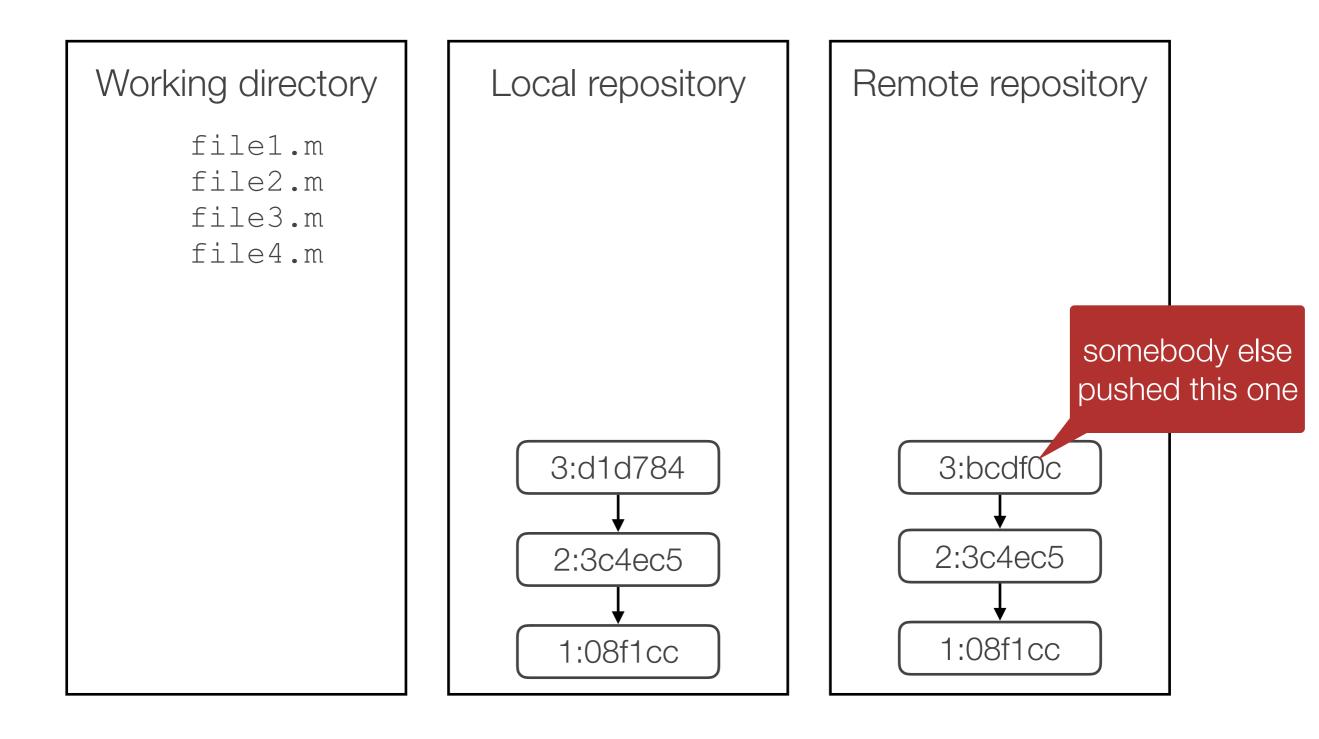


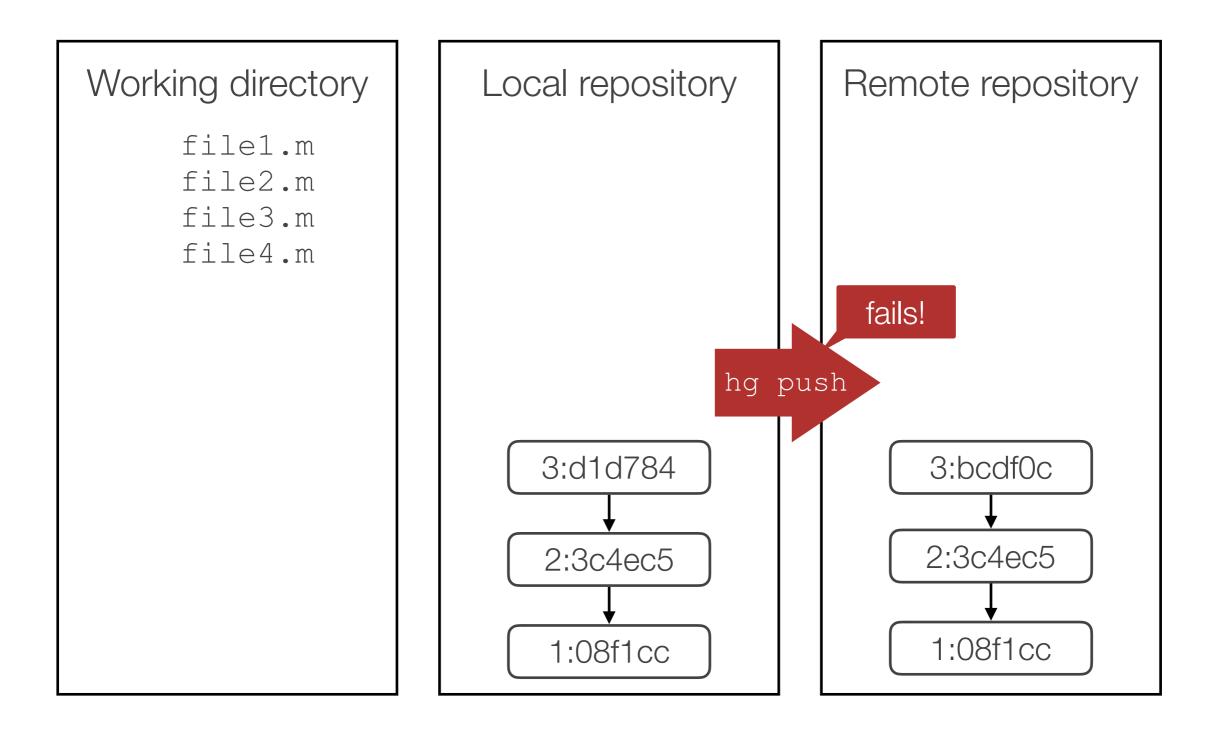


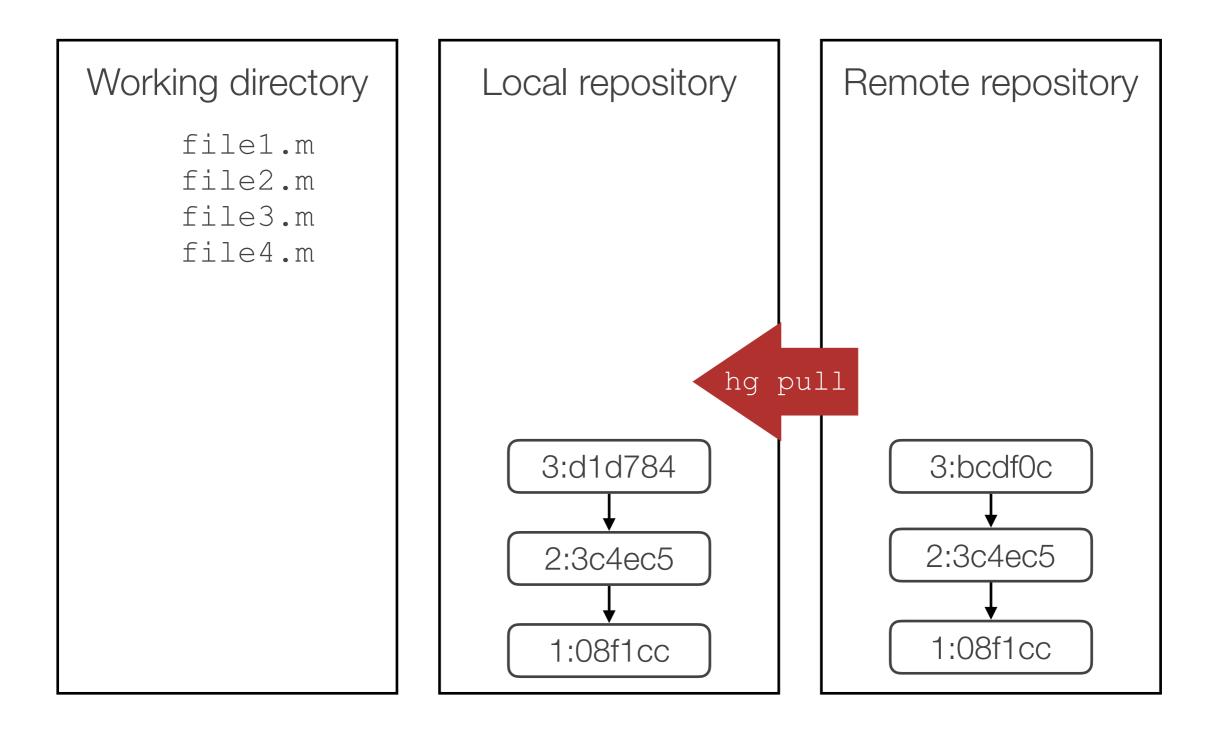


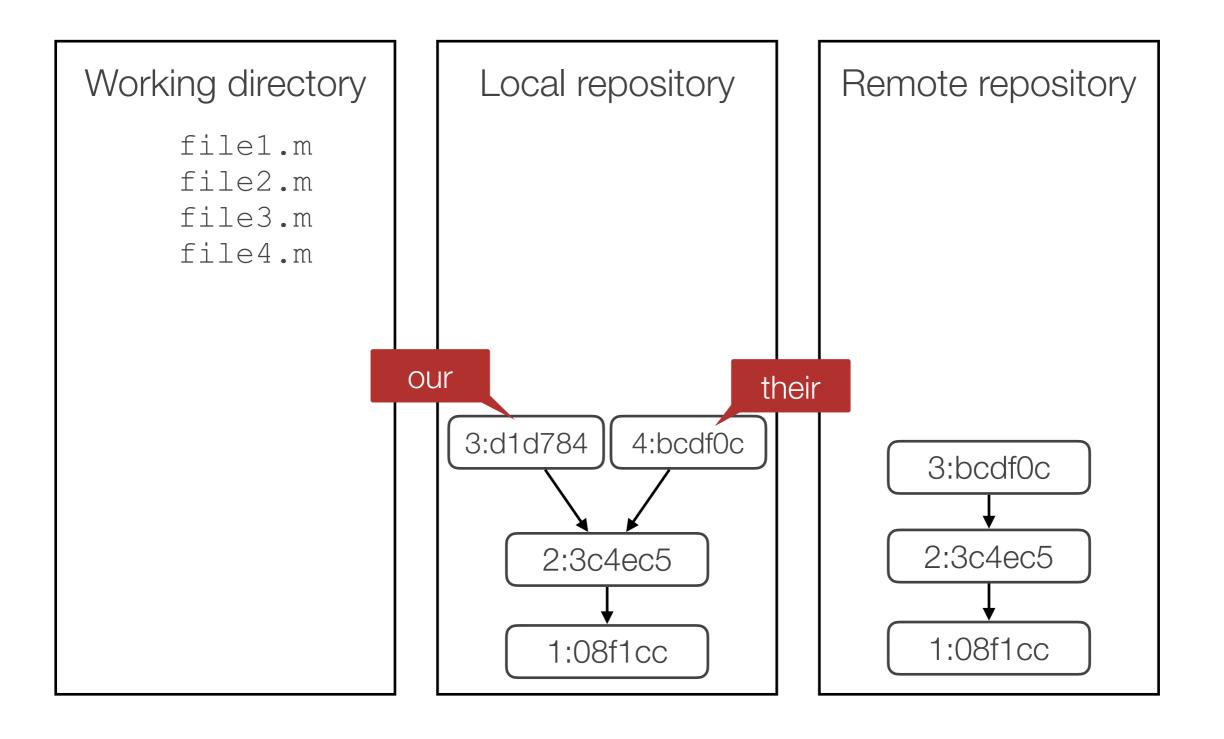


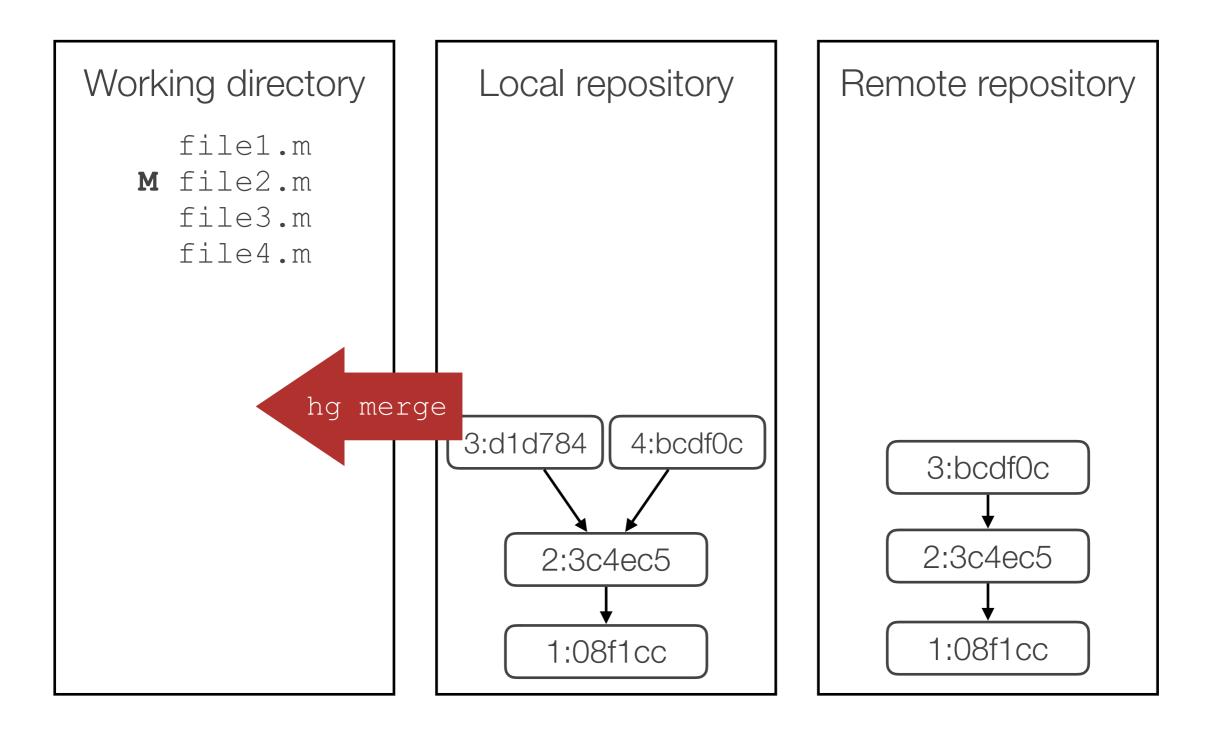


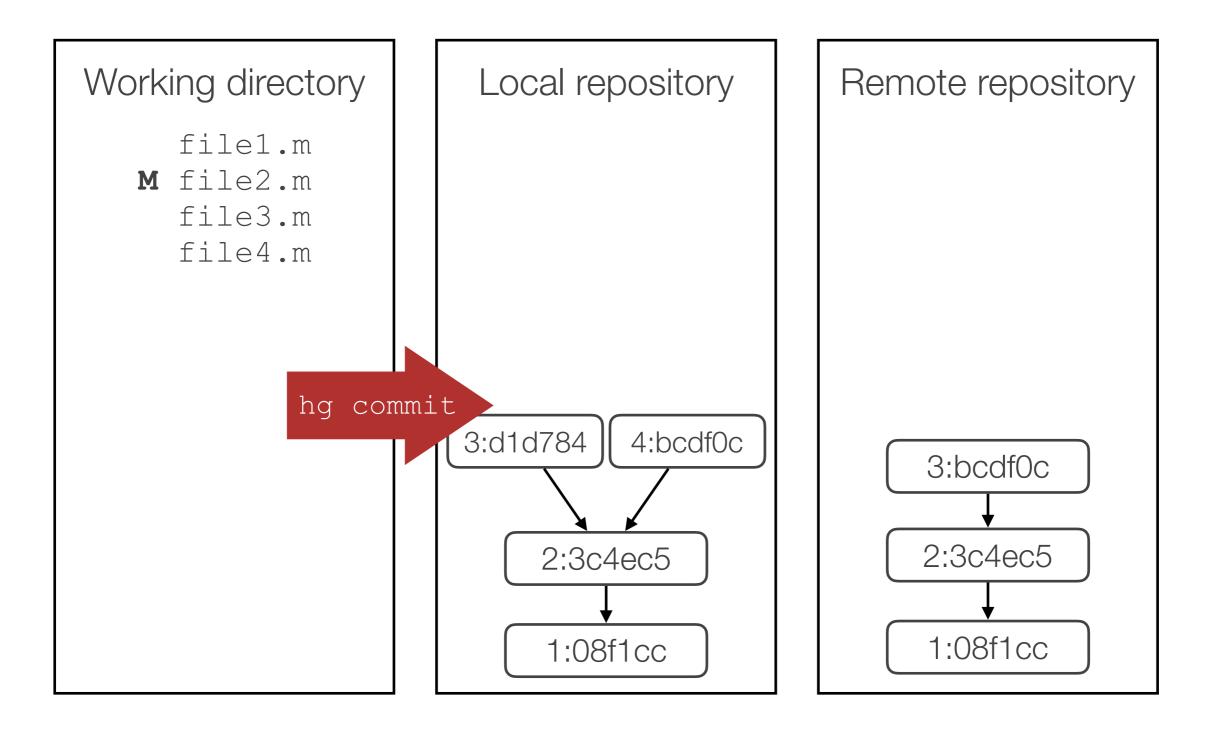


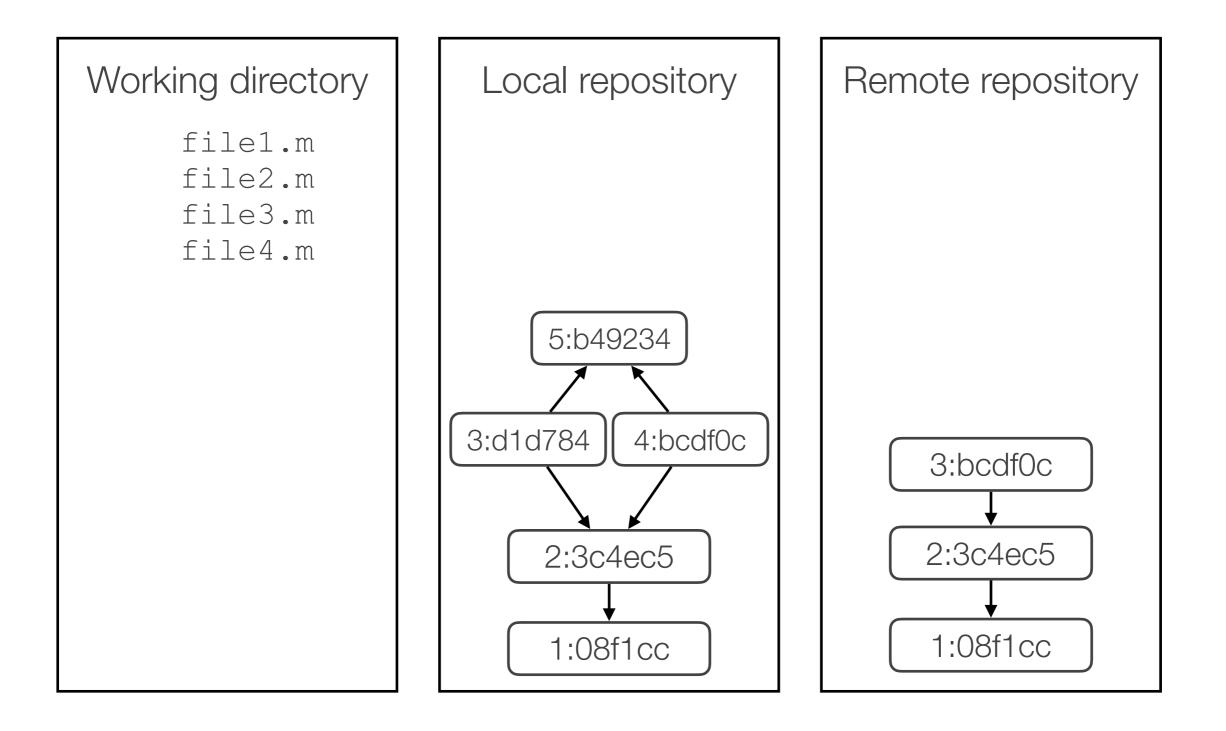


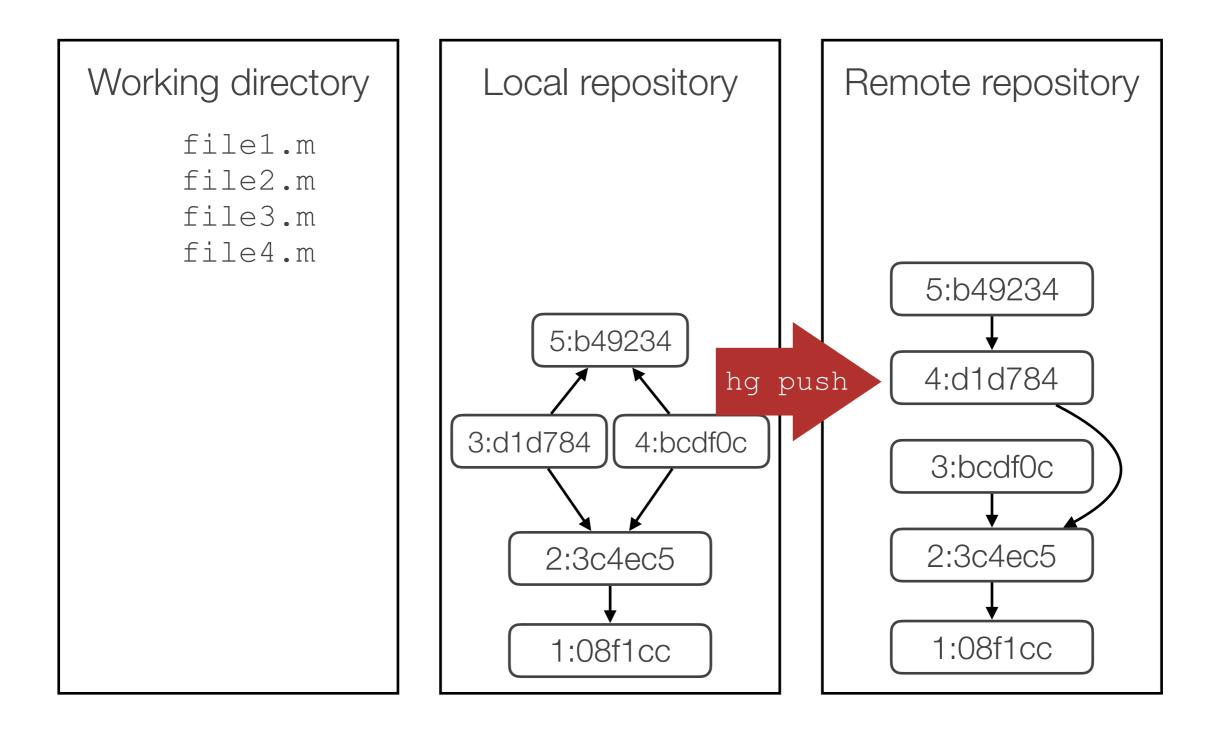


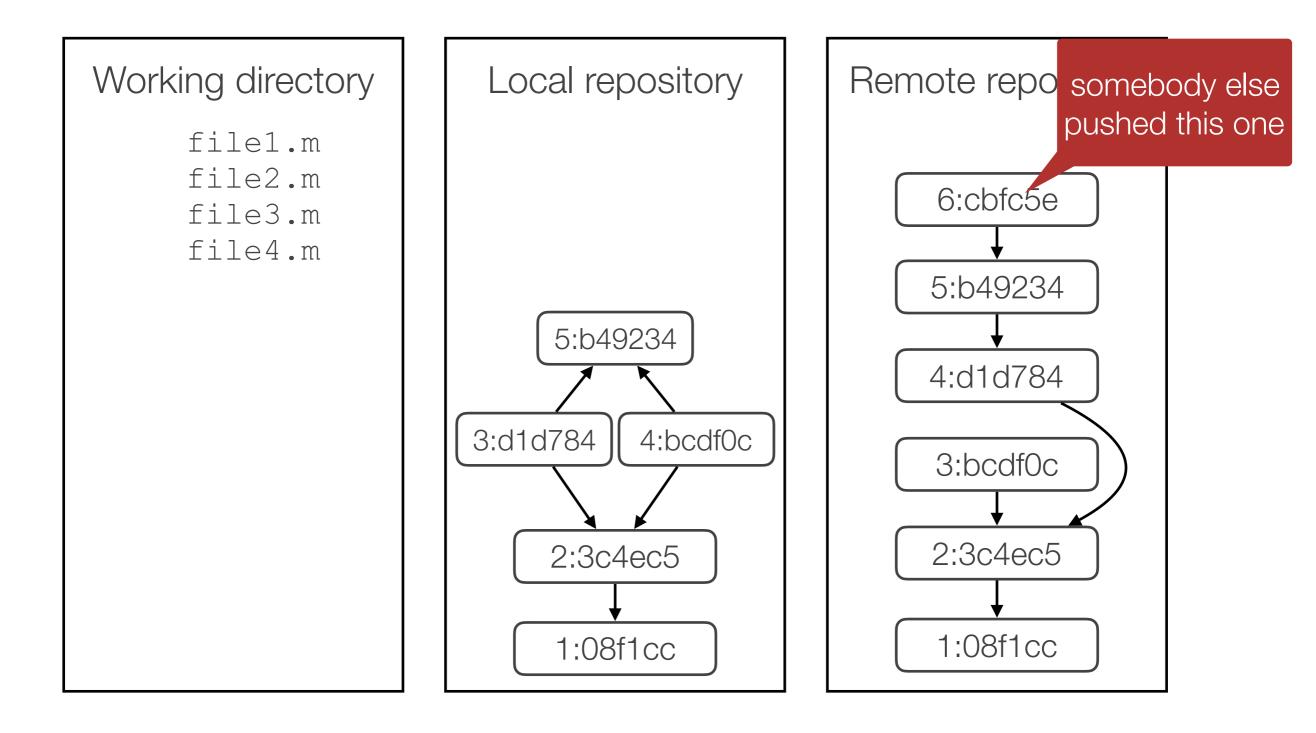


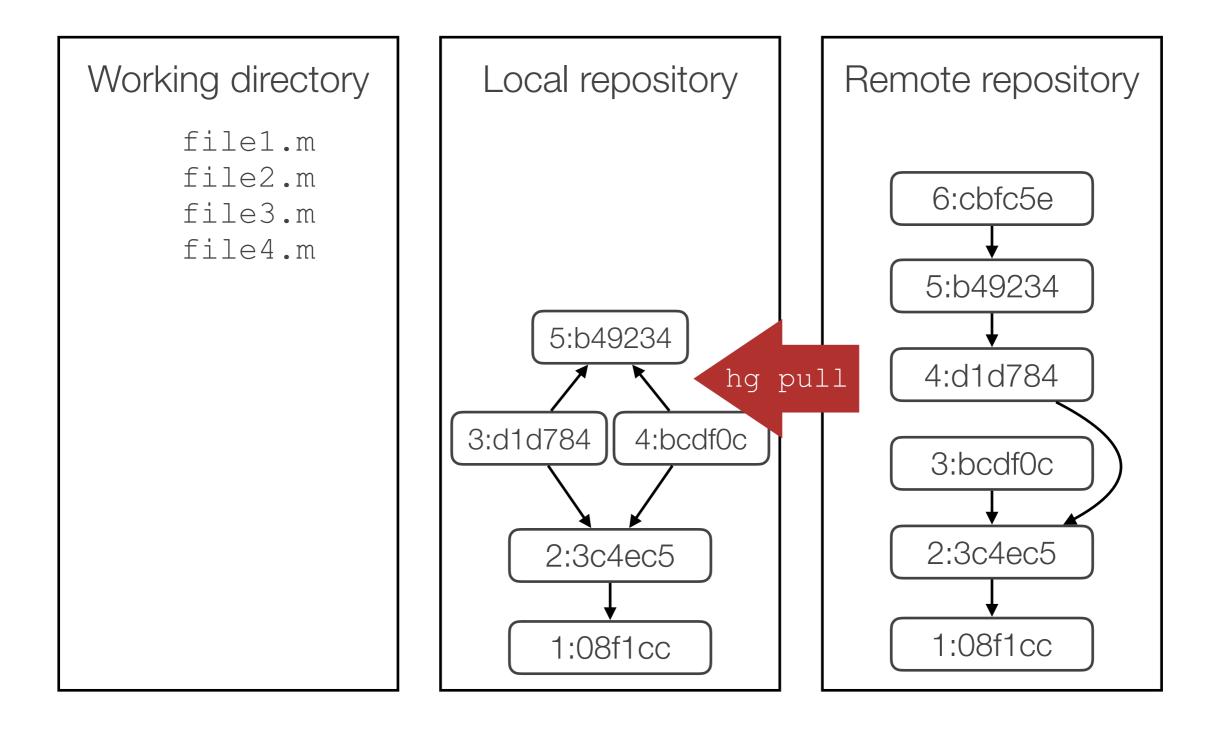


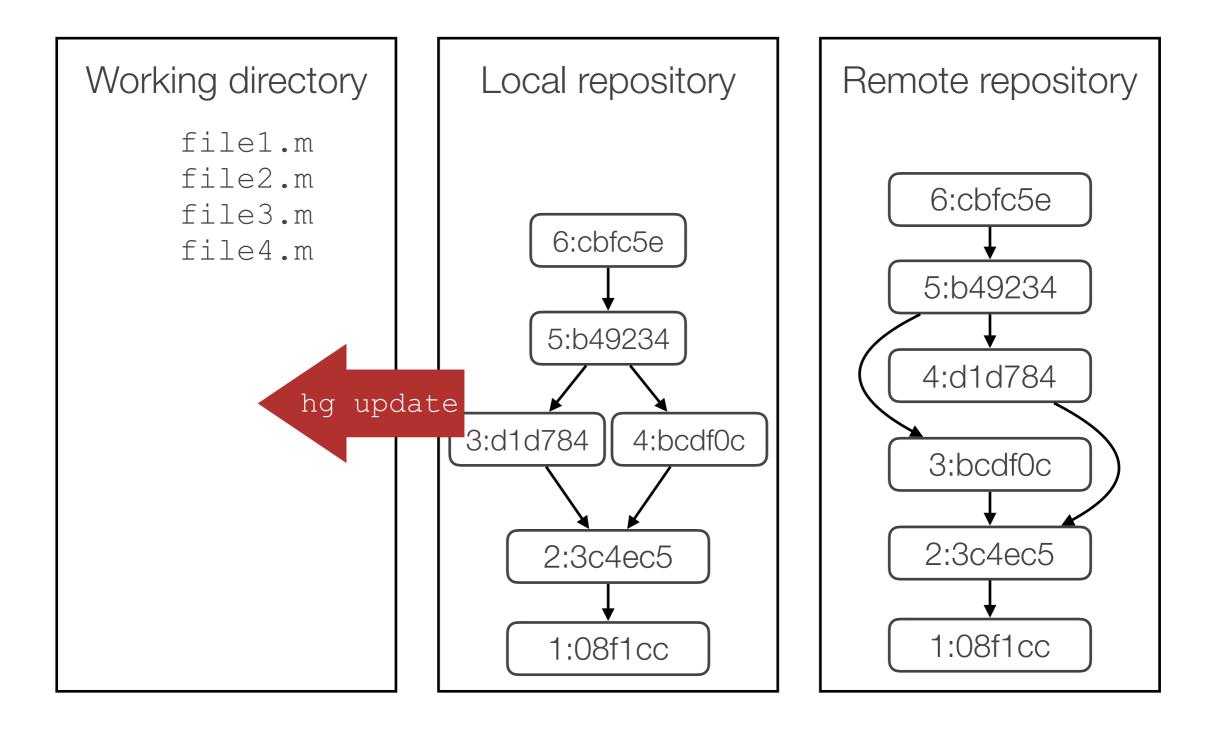








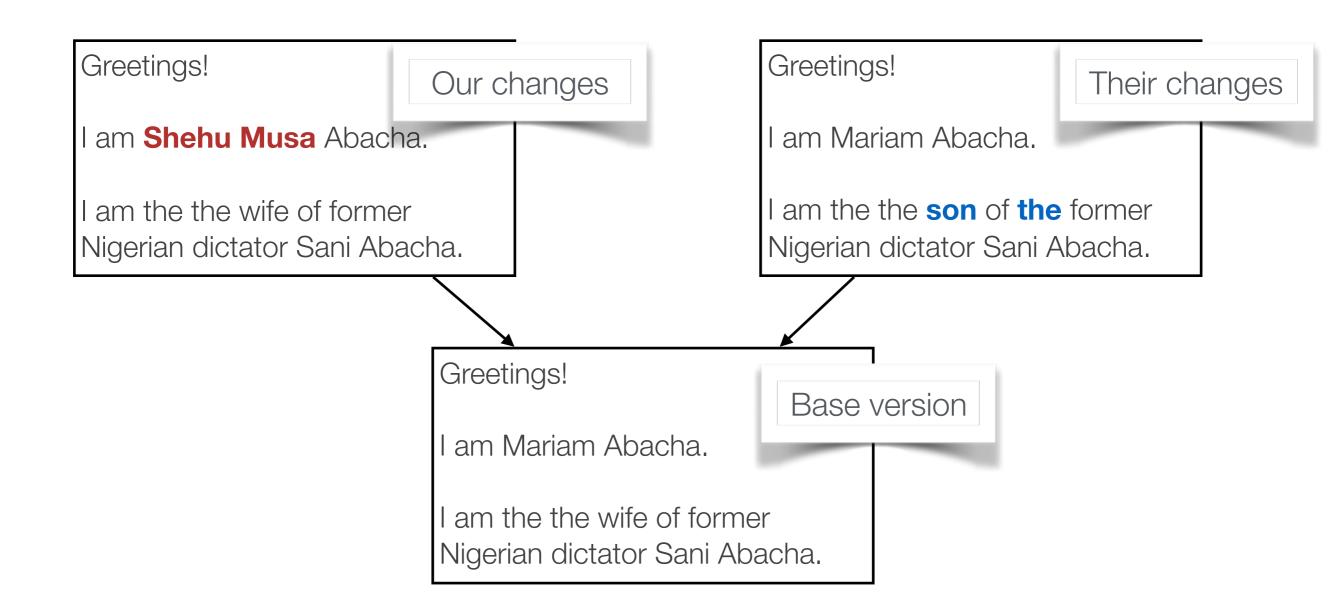




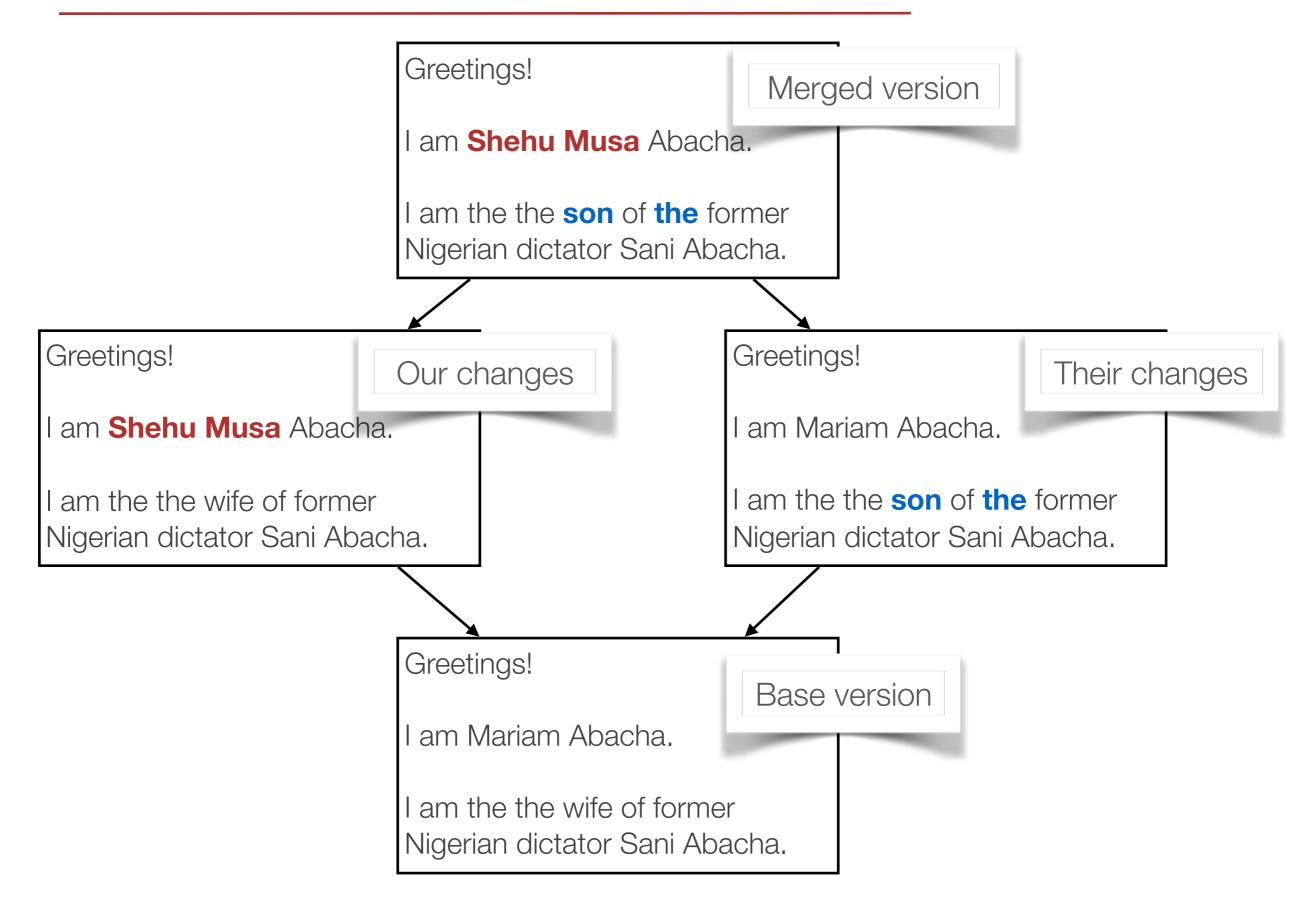
Sometime Merging is Fully Automatic...

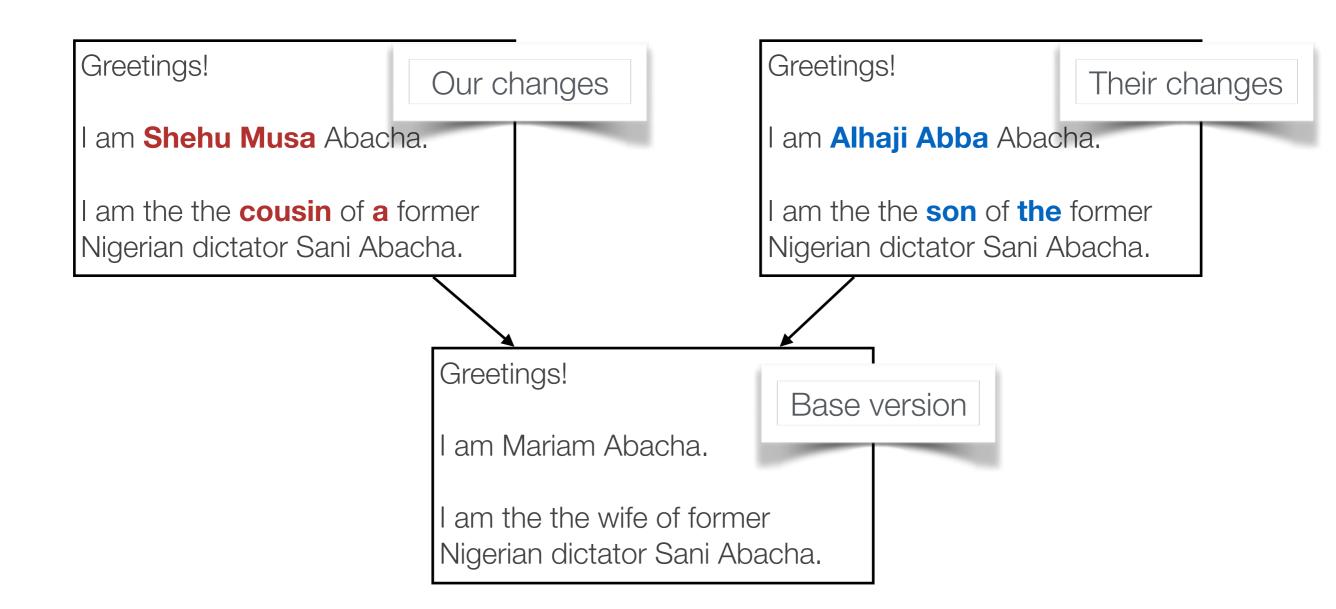
Greetings! I am Mariam Abacha. I am the the wife of former Nigerian dictator Sani Abacha.

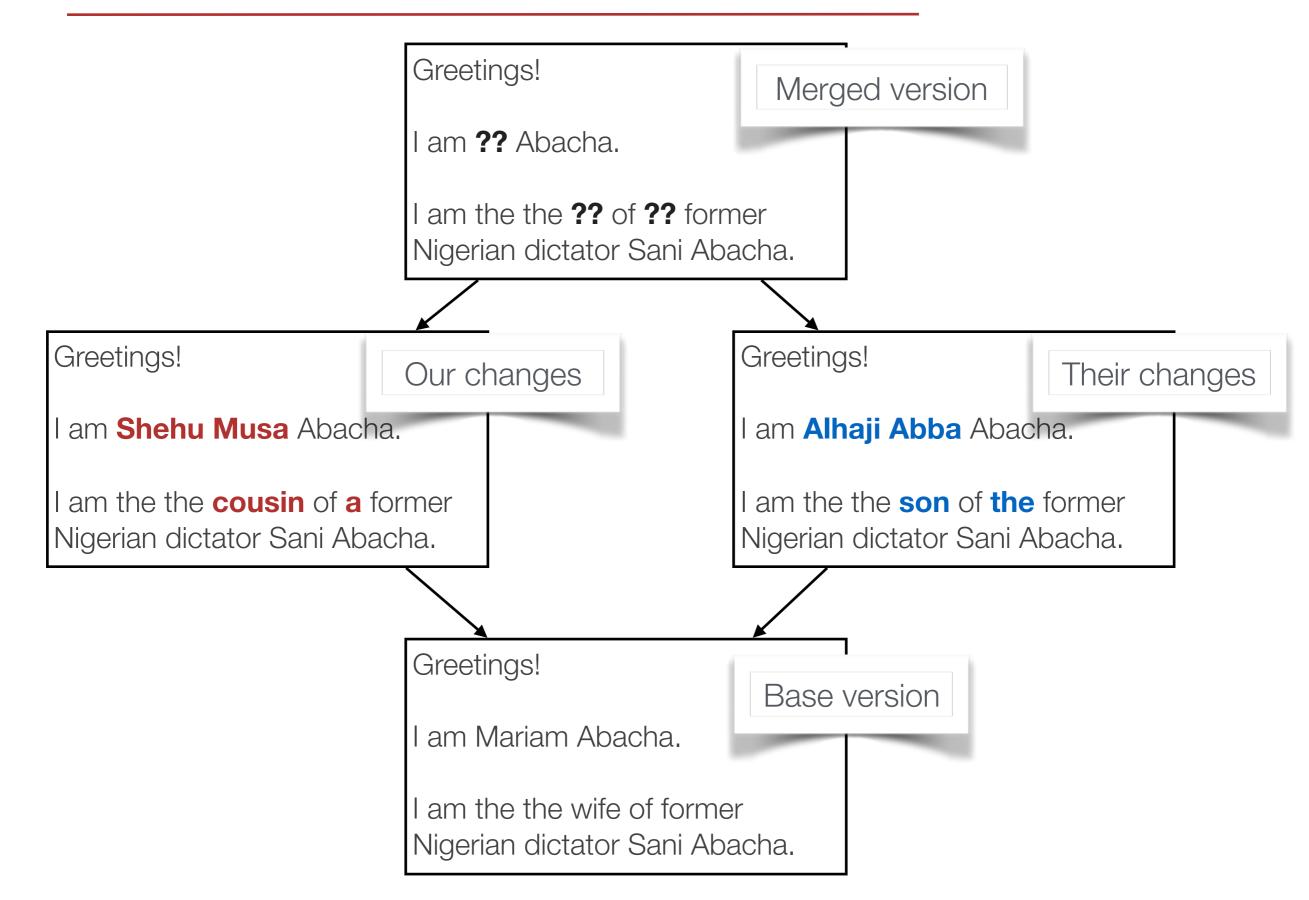
Sometime Merging is Fully Automatic...

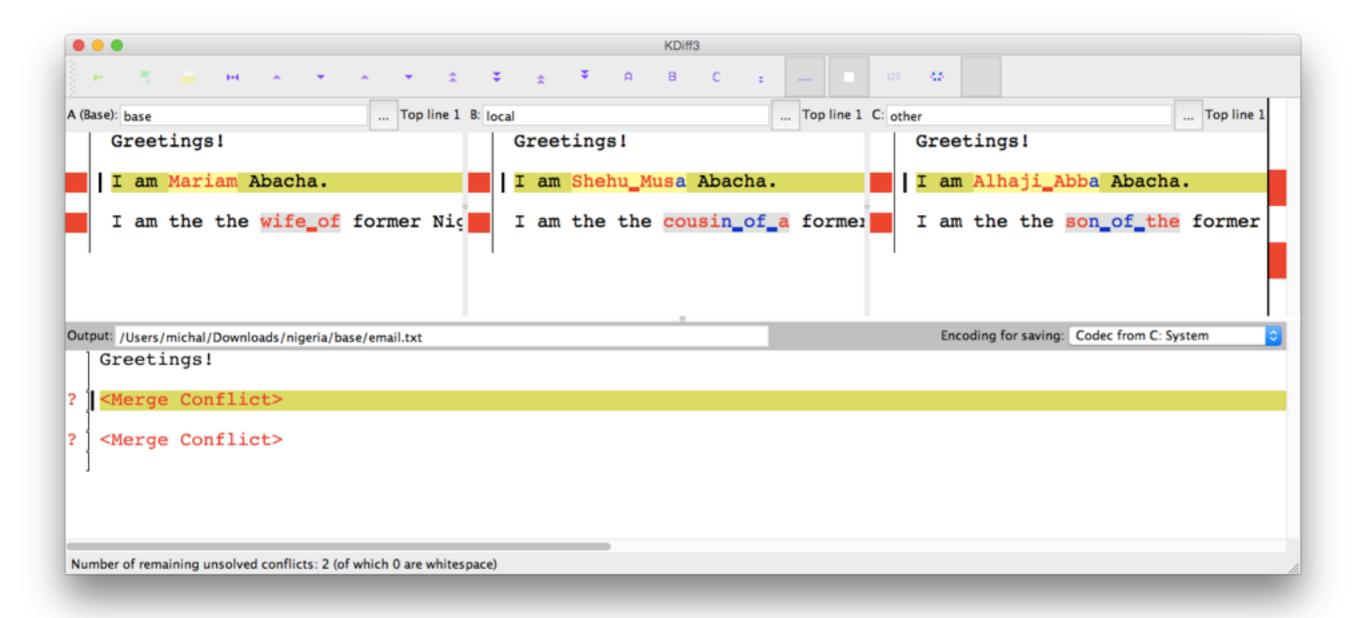


Sometime Merging is Fully Automatic...

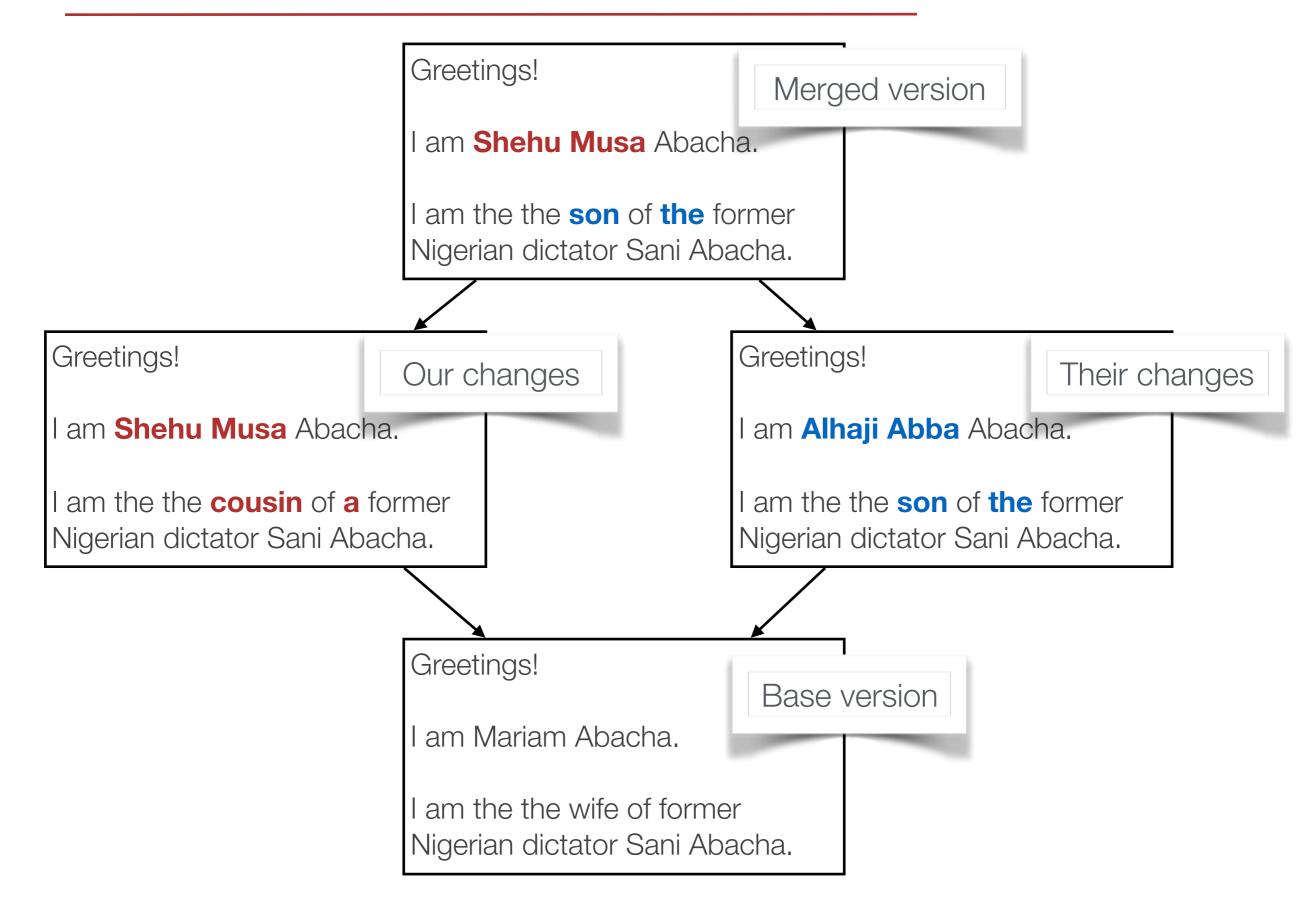


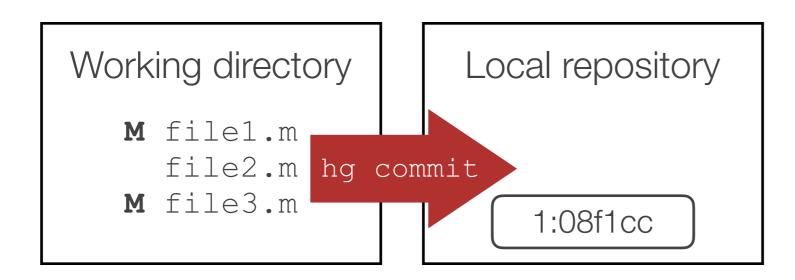






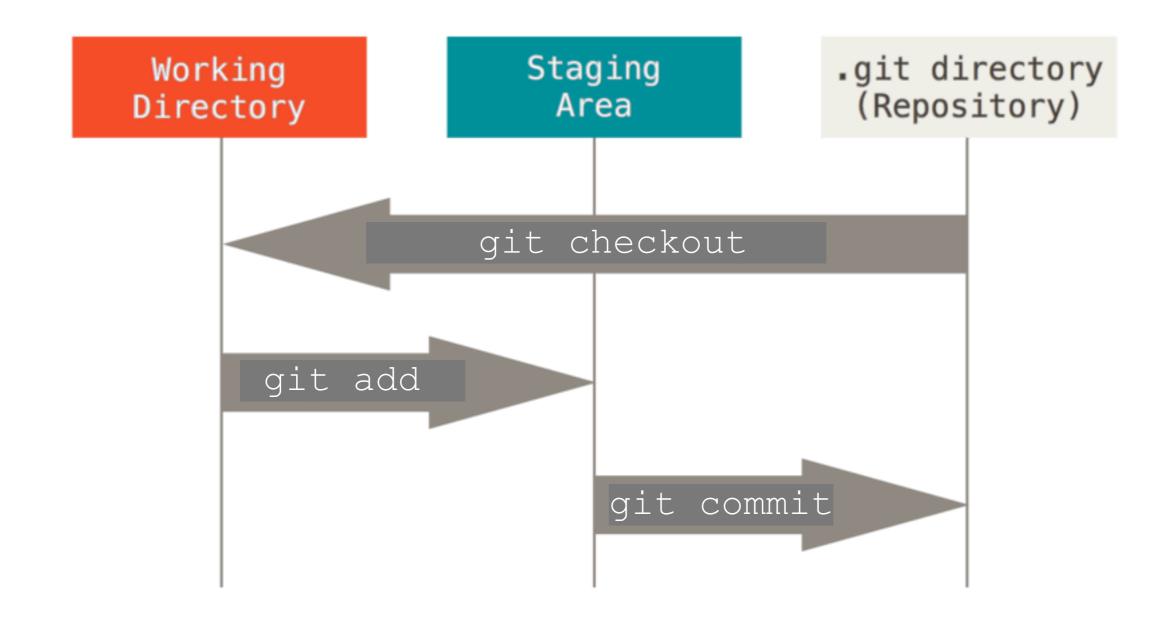
● ● ● KDiff3						
₩ <u>₩</u> ₩ ★ ¥ ★ ¥ ± ¥ ± ¥ A B (2 z 123 44					
A (Base): base Top line 1 B: local	Top line 1 C: other Top line 1					
Greetings! Greetings!	Greetings!					
I am Mariam Abacha. I am Shehu_Musa Ab	Jam Alhaji_Abba Abacha.					
I am the the wife_of former Nic I am the the cousi	n_of_a former I am the the son_of_the former					
Output: /Users/michal/Downloads/nigeria/base/email.txt	[Modified] Encoding for saving: Codec from C: System					
Greetings!						
B] I am Shehu Musa Abacha.						
C I am the the son of the former Nigerian dictator Sani Abacha.						
j						
Number of remaining unsolved conflicts: 0 (of which 0 are whitespace)						



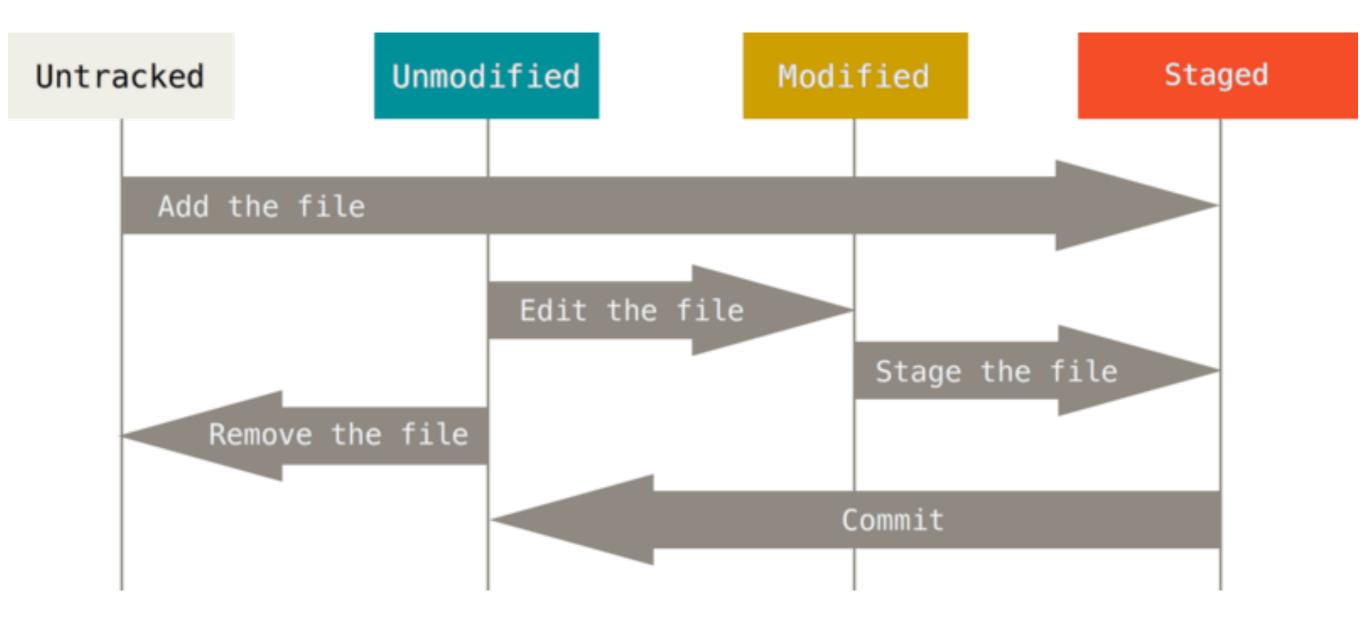


In Mercurial, hg commit automatically commits all modified files In Git, things a bit more complicated...

Sidenote: Committing in Git



Sidenote: Committing in Git



Learn the command-line syntax first to understand concepts

Then use GUIs:

- my choice: SourceTree (hg, git, win, mac, free)
- other options: GitHub Desktop, GitEye, SmartGit, ...

V O Q (• • • • • • •	a 1 a	2	r 🍆 👘	5	ห	> _	- Q	
View Co	mit Update Revert Shelve Add Remove	Add/Remove Pull Push	Branch M	erge Tag	Show in Finder	Hg Flow	Terminal	Setting	
FILE STATUS	All Branches						Jump to:	4	
BRANCHES	Description Uncommitted changes					Revision	Author	Date 19 Nov 2016 13:4	
₽ derault ₽ develop	Image: tip Image:					828 827	Michal Kvasnica Michal Kvasnica	17 Nov 2016 08:3 17 Nov 2016 02:5	
v 📄 hotfix						826 825	Michal Kvasnica Michal Kvasnica	16 Nov 2016 09:4 15 Nov 2016 15:5	
r ieature P enumpap						824 823 822	Michal Kvasnica Michal Kvasnica Michal Kvasnica	15 Nov 2016 15:2 15 Nov 2016 15:2 15 Nov 2016 15:2	
BOOKMARKS					821	Michal Kvasnica Q, Search	15 Nov 2016 12:1		
REMOTES	CHANGELOG CHANGELOG				CHANGELOG				
Ø default									
SHELVED	<pre>source/mpt/tests/ui/pwasystem/test_pwasystem_reachset_07_pass.m PWASystem/reachableSet: correctly deal with intermediate empty sets</pre>			File contents 1 + MPT 3.1.6				Reverse hunk	
SUBREPOSITORIES				2 + ====== 3 + 4 + Fixed bugs 5 +					
	Revision: 828		6 + *	PWASystem/reach	ableSet() now co	rrectly d	eals with intermedi	ate empty sets	



Demo

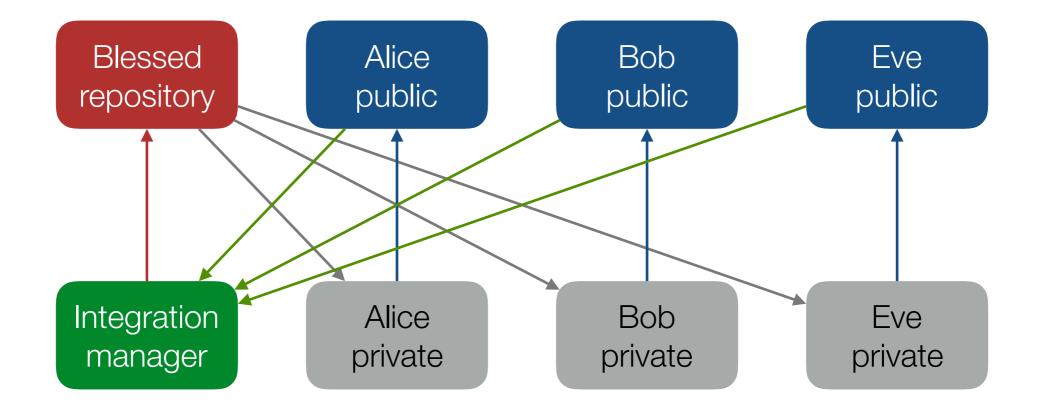
Public Repository Hosting

Cloud-based:

- Bitbucket (Mercurial+Git)
- GitHub (only Git)
- GitLab (only Git)

Self-hosted:

- Rhodecode (Mercurial+Git)
- Bitbucket Server (only Git)
- GitLab (only Git)



Bitbucket vs GitHub

Both:

- are free for academia (Bitbucket even free for small teams)
- offer code hosting, code review, issue tracking, wikis
- support pull requests

My choice: Bitbucket

But GitHub is more popular and visually more appealing



Demo

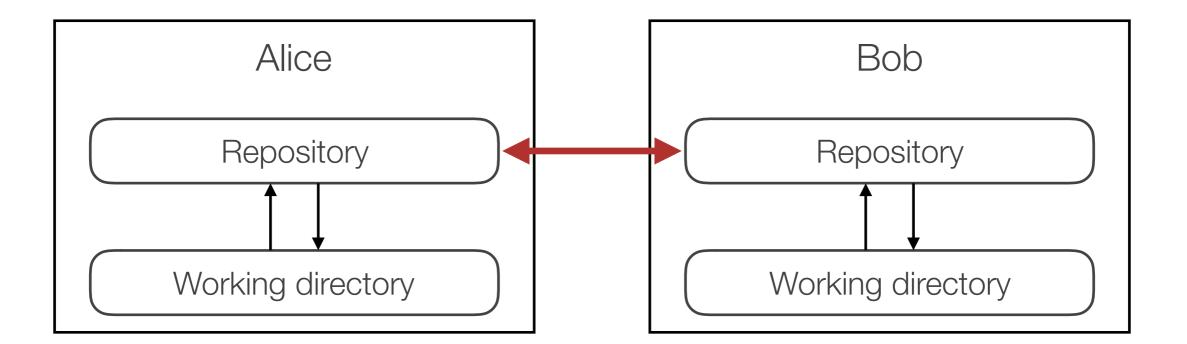
Pull Requests

Main idea: everybody has a local copy of the repository

- every developer can commit/update at any time to his/her repository
- no permanent connection required

Question: communication? ("Hey, I made a change, you need to update")

Answer: pull requests



Demo

Mercurial vs Git: Summary

DOs:

- learn the command line syntax to understand the concepts
- install a merging tool (e.g. kdiff3)
- pick a GUI that supports both systems
- write meaningful commit messages
- use pull requests

DONTs:

- do not lead religious wars about your choices, write code

Agenda

Today

- version control systems and collaborative development
- Mercurial, Git, Bitbucket, GitHub
- providing support

Tomorrow

- unit testing
- documentation
- dissemination

Providing Support

Request tracking:

- email + a request tracking system

Bug tracking:

- Bitbucket/GitHub

Discussion groups/fora:

- Google Groups

Create a generic email (e.g. mpt@) instead of using username@

- allows to dispatch the email to a support group
- lets you to change your affiliation later

Keep track of which emails have been answered:

- sounds easy, but difficult if we have multiple persons providing support
- use Trello as a simple request tracking system
- or set up a more complex system (Request Tracker, Roundup, ...)

Trello

🖸 Boards 🔎	🖬 Trello	+ Michal Kvasnica 🕡 🗘				
MPT 🏠 🛆 Private		··· Show Menu				
Reports ···	Bugs ···	Features				
Re: Issue with missing regions using the parametric QP solver Oct 28 ■	Wrong volume due to failed vertex enumeration	add mpt_enum_pqp				
Re: How can i define polytopic uncertaint system in mpt 3 and	consistent handling of Opt/yalmip, Opt/removeEqualities ≣ Add a card	introduce Opt/getRegion ≡				
compute its invariant set ? Oct 28 ≡		unified progress reports ≡				
mpt+cplex ② Oct 30 ≡	Add d Gard	consistent handling of MPTOPTIONS.verbose				
Strange cddmex error for vertex-ray polyhedron		 Drop "_pass"/"_fail" suffixes from tests				
PWA reachable set computation problem ② Nov 17 ■		PolyUnion methods should not require the function name if the object has a single function				
Add a card		mpt_voronoi: take points stored row-wise				
		mptopt: allow mptopt('modules.ui.invariantSet.max Add a card				

Bug Tracking

10 years ago:

- many options: Bugzilla, Redmine, Trac, Roundup
- tried them all, liked Roundup the most

Today:

- use issue tracking modules in Bitbucket and GitHub

Persuade your users to submit bug reports

How to Write Good Bug Reports

Check if you are using the latest version and whether the bug is known

File each issue separately

- bad: "fun1 errors, fun2 misbehaves, and, btw, there is a typo in fun3"

Give the report a descriptive name

- bad: "IT DOES NOT WORK!"
- good: "fun1 breaks when called with a string"

State what is the expected result and what is the actual behavior

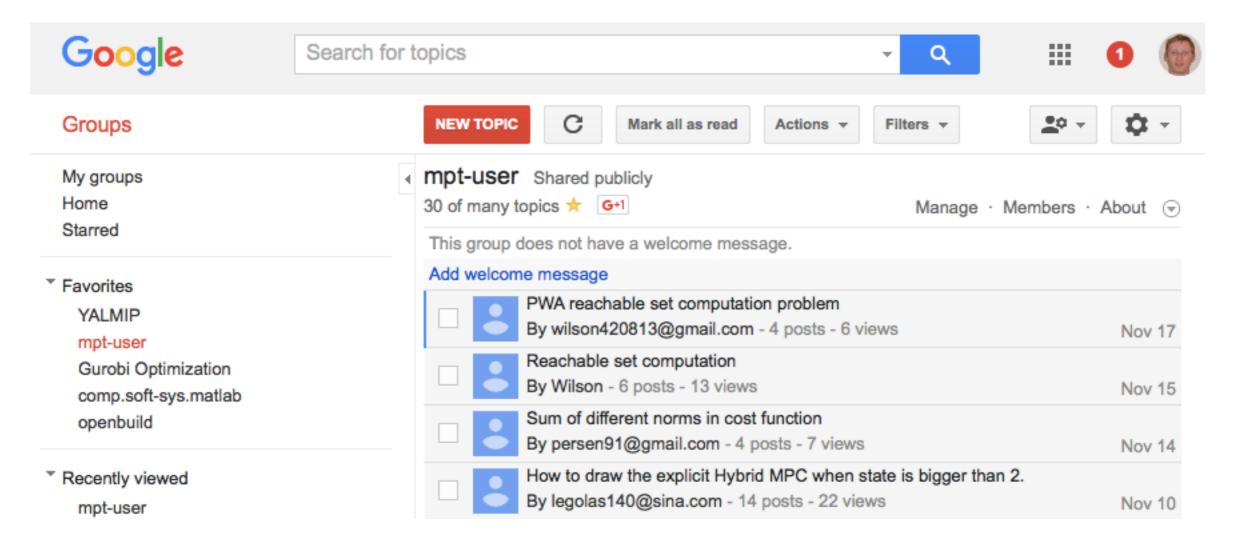
Give steps to reproduce the issue

- give information about your setup (OS/Matlab version, etc.)
- provide a minimal code!



Fact: same questions are asked over and over again

Idea: let users provide support to each other



Side note: redirect the support mail to the group



Demo

Agenda

Today

- version control systems and collaborative development
- Mercurial, Git, Bitbucket, GitHub
- providing support

Tomorrow

- unit testing
- documentation
- dissemination