2nd GouTP @ SCEE

Open-sourcing your code with git

Lilian Besson

SCEE Team, IETR, CentraleSupélec, Rennes

Thursday 9th of November, 2017

2nd 2017/18 GouTP @ SCEE

- *About:*
  - Version control with git
  - Share your simulation code for reproductibility
  - Open-source your code with git
- *Date:* 9th of November 2017
- *Who:* Lilian Besson

What's a "*GouTP*" ?

- **Internal monthly technical training session**
- Usually: *Thursday 3pm - 3:30pm*
- With coffee and sweets : we relax while training !

Initiative of Quentin and Vincent in last January... Continued by Rémi, Rami, Muhammad and Lilian !

Not only @ SCEE ?

- 2nd and 3rd GouTP will be open to the *FAST team* $\rightarrow$ If success, next ones will be open to other research teams @ Supelec Rennes

Agenda for today [30 min]

1. Concept of version control with `git` (demo) [10 min]
2. Research collaboration on code or articles with `git` (examples, good practice) [5 min]
3. Why we should all share our simulation code online, and under an open-source licence (and even LaTeX) [10 min]
4. Example of open-sourcing the MATLAB code and LaTeX code from a recent article [5 min]

Why Git ?

Version control . . .

- Veru useful to:
- Never lose your code
- Keep track of progress, revert changes when needed
- Collaborate easily and asynchronously
- Git is used everywhere , easy to learn and powerful
- Free online hosting: Bitbucket, GitHub, GitLab etc. . .

Tutorial online !

- Try this please → `Try.GitHub.io`

Quick *live demo* of Git

Meta demo ?

- I store all my slides on GitHub…
- … Including the source code for this one
- Let see together ! → *live demo* of local use of `git`
- basic commands for a use in a terminal
- or in a graphical interface (e.g., inside your IDE)

→ *live demo* of the online repository (on GitHub)

Research collaboration with git

Why ?

- Easiest way to collaborate on code or article
- No email, no painful Dropbox/Drive synchronization ...
- Full control on your files' history and the synchronization !

How ?

1. Create a (*private*) repository that your colleagues can access
2. Where? *Example*: Bitbucket, GitHub (with student pack), GForge @ Inria, OverLeaf (for LaTeX)...
3. Start collaborating with no sweat !

Share your simulation code online

Why ?

- Everyone can (hopefully) reproduce your code and results
- Show to the world that you do *serious reproducible* science !!

How ?

1. Clean up your source code, and add a few comments
2. Write a small `README.md` file to explain: how to run your code, for which article it was used, conditions of usage etc
3. Maybe add an example, or figures / screenshots
4. Ex: `Bitbucket.org/SCEE_IETR/Testbed_Monitor` for an internal tool, or `Bitbucket.org/SCEE_IETR/RL_Slotted_IoT_Networks` for an article

Join the open-source community !

- `ChooseALicense.com` to pick a license suiting your needs
- By default HAL uses a **Creative Commons** license (with various flavors). Example : `HAL.Inria.fr/HAL-01575419`
- But arXiv does not specify the license (on document and source) : that's bad ! No one can use your code if you do not specify any copyright or usage conditions...

My advice ?

- I suggest the **MIT License** for simulation code (short & well-known) and **Creative Commons** for documents and LaTeX

Example of sharing on Bitbucket the simulation code from an article

It takes 10 minutes:

1. Clean up the MATLAB files
2. Add a few comments in the tricky parts
3. Add a header to the files stating the copyright
4. Choose a license and add a `LICENSE` file
5. Write a `README.md` file in the folder
6. Create the repository, `git add` all the files
7. `git push`, check the result, and relax !

→ `Bitbucket.org/SCEE_IETR/RL_Slotted_IoT_Networks`

And open-sourcing the LaTeX code?

Note: this is not against the copyright policies of conferences/journals if you do not share the PDF...

- Not so useful for articles with basic templates, but why not?
- Can help your colleagues if you use a nice template for posters or slides
- Can also help when writing your thesis, you can copy-paste equations from your colleagues' articles instead of re-writing...
- Example: `Bitbucket.org/LBesson/Multi-Armed-Bandit-Learning-in-IoT-Ne`

Conclusion

- I hope you got an overview of how to use `git`
- Why it can be a good idea to share your simulation code
- And why choosing an open-source license is smart !

Your mission, if you accept it...

- *Padawan level:* Train yourself on git → `Try.GitHub.io`
- *Jedi level:* Release some simulation code online !
- *Master level:* Release *all* your code (and LaTeX) online !!

Thanks for joining ! Contact us if you want to do a GouTP!