INTRODUCTION TO NOTEBOOKS AND JUPYTER

GOOD PRACTICES AND COMPUTATIONAL REPRODUCIBILITY

Amandine Perrin, Etienne Kornobis, Bertrand Néron, Frédéric Lemoine
2022/05/30
Institut Pasteur

PART 1

Introduction

Institut Pasteur

1.1 The concept of literate programming

Donald Knuth

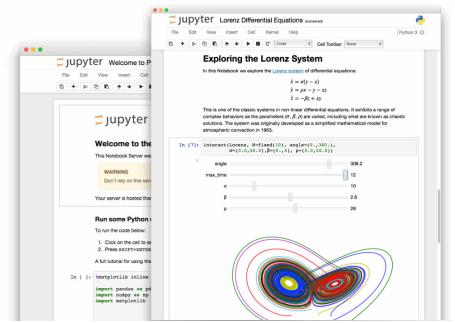
Literate programming: Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do.

1.2 Notebooks

The concept of notebook

A single place where live happily together:

- code
- documentation
- reasoning
- visualizations



jupyter Welcome to P
File Edit View Insert Cell
jupyter

Welcome to the

This Notebook Server wa

WARNING
Don't rely on this ser

Your server is hosted thar

Run some Python

To run the code below:

1. Click on the cell to se
2. Press SHIFT+ENTER

A full tutorial for using the

```
In [ ]: %matplotlib inline

        import pandas as pd
        import numpy as np
        import matplotlib
```

jupyter Lorenz Differential Equations (autosaved)
File Edit View Insert Cell Kernel Help
Python 3
Code
Cell Toolbar: None

Exploring the Lorenz System

In this Notebook we explore the Lorenz system of differential equations:

$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = \rho x - y - xz$$
$$\dot{z} = -\beta z + xy$$

This is one of the classic systems in non-linear differential equations. It exhibits a range of complex behaviors as the parameters (σ, β, ρ) are varied, including what are known as *chaotic solutions*. The system was originally developed as a simplified mathematical model for atmospheric convection in 1963.

```
In [7]: interact(Lorenz, N=fixed(10), angle=(0.,360.),
                 σ=(0.0,50.0),β=(0.,5), ρ=(0.0,50.0))
```

angle	308.2
max_time	12
σ	10
β	2.6
ρ	28

1.2 Notebooks

Many flavors of notebooks

Many technologies have been developed in this direction:

- Jupyter: `https://jupyter.org/`
- R Markdown: `https://rmarkdown.rstudio.com/`
- Apache zeppelin: `https://zeppelin.apache.org/`
- Google Colaboratory: `https://colab.research.google.com`
- Observable (client-side): `https://observablehq.com/`
- Spark notebooks: `http://spark-notebook.io/`
- Beaker (engulfed by Jupyter)
- ...

1.2 Notebooks

Caveats

Notebooks downsides:

- Can be cumbersome (cell order execution...), not necessarily good as first entry point in Programming
- Beware of bad coding practices (no proper module/library design)
- Difficult for source control
- Testability

But these issues are currently addressed !

1.3 Jupyter

Why choosing Jupyter ?

- More than 40 languages supported
- Extensibility
- Community
- Recognized and pretty formatted by GitLab and GitHub

1.3 Jupyter

The main idea of Jupyter

An open-source web application to interactively represent a single json file containing:

- code
- documentation
- reasoning
- visualizations (simple and interactive)

Features overview:

- Easy sharing
- Multitude of export: interactive notebook, voilà dashboard, html blog, presentation, or simple script.
- the future of scientific article publication (at least in programming related fields)

1.3 Jupyter

Many IDE flavors for Jupyter notebooks

- Jupyter notebook (+ extensions)
- Jupyter lab (+ extensions)
- nteract
- Your own editor (visual studio code / pycharm / jupytext / Emacs modes ...)

JupyterLab is the main IDE for Jupyter now but:

- Extensions do not necessarily work for both Jupyter notebook and lab
- Nice features in nteract but cannot use previous extensions

PART 2

Setup

Institut Pasteur

2.1 Installation

Documentation: https://jupyter.org/install

Locally, Using conda (recommended on jupyter website):

```
conda env create -n jupyter jupyterlab
```

If necessary extension can be installed from the JupyterLab interface.

On maestro:
See `https://confluence.pasteur.fr/display/FAQA/How+to+use+Jupyter-Notebook+on+the+cluster`

2.2 Exporting / Sharing

Nbconvert

Jupyter files can be converted in many formats using *nbconvert*:

- HTML
- Reveal JS slides
- LaTeX
- PDF
- Markdown (md)
- ReStructured Text (rst)
- executable script

2.2 Exporting / Sharing

Helpful services

A lot of services are provided to ease notebooks sharing:

- JupyterHub: `https://jupyter.org/hub`
- Binder: `https://gke.mybinder.org/`
- Voilà app: `https://voila.readthedocs.io/en/stable/using.html`
- Github/Gitlab formatting.
- Usefull extensions e.g. Jupytext: `https://jupytext.readthedocs.io/en/latest/install.html`

PART 3

Demonstration

Institut Pasteur

Overview

- Mixing Code and markdown
- Visualizations
- Interactivity
- Virtual environment support: conda environments
- Presentations (Reveal.js and Rise)