

# Assignment 2 Setup

Assignment 2 involves using two software packages, [PyMOL](#) and [PyRosetta 3](#). Both packages, as well as Python 2.7, are already installed on the VPTL Macs (all the cluster computers in the libraries and dorms) as well as the FarmShare `rye` machines. Below we give instructions on how to install these packages locally on your own machine, how to access them on the VPTL, and how to access them remotely on `rye`.

## Suggested workflow

### *OS X & Linux Users*

We recommend that OS X and Linux users first try installing the software on their own machine. If you run into problems, we next recommend you work on the VPTL Macs. If you do not have access to the VPTL Macs, you can interface with the software remotely on the `rye` machines.

### *Windows Users*

Installing and getting PyRosetta to work on Windows is tricky, so we recommend that Windows users work on the VPTL Macs. If you do not have access to these machines, we recommend working remotely on the `rye` machines. If you're having trouble setting up, do not hesitate to come to the TAs' office hours or posting to Piazza.

## Working on your own OS X/Linux machine

If you have not already downloaded and installed PyMOL, go ahead and do that now. The installation is very straightforward. Instructions for PyRosetta 3 installation are below.

Detailed notes on PyRosetta installation from <http://www.pyrosetta.org/dow>

1. Obtain a [Rosetta license](#) from to receive a username and password.
2. Download **PyRosetta 3** (you will run into import problems with PyRosetta 4 ). Link here: <http://www.pyrosetta.org/dow/pyrosetta3-download>. Make sure to download **release.monolith**
3. Unpack the downloaded file to the location of your choice to create the PyRosetta directory.  
(From a terminal/console window, you can unpack the archive using the command:

```
tar -vjxf PyRosetta-<version>.tar.bz2
```

Please note, there is no special install procedure required; after unpacking, PyRosetta is ready to use. So unpack it to the location from where you want to execute it.

```
$ cd ~/Downloads # navigate to the directory which you've downloaded the tar.bz2
$ tar -vjxf PyRosetta-<version>.tar.bz2 # Unpack the tar.bz2
$ cd PyRosetta-<version> # navigate to the unpacked pyrosetta directory
```

4. To set up the PyRosetta library file paths, you must execute the following command from within the new PyRosetta directory:

```
source SetPyRosettaEnvironment.sh
```

You will have to do this step every time you use PyRosetta. Alternatively, you can edit your bash profile file to have it run automatically every time you open up a new terminal window:

```
# Alternative Method
$ vim ~/.bash profile # It's ~/.bash.profile for Mac users and ~/.bashrc for linux users
# press 'i' and type:
$ source SetPyRosettaEnvironment.sh
# press 'esc', ':wq', then 'Enter'
$ source ~/.bash profile # applied every time you open a new terminal
```

5. Start Python.

```
$ python
```

6. In Python, you should be able to import the PyRosetta library with the commands:

```
$ >>> import rosetta
$ >>> rosetta.init()
```

7. If this step does not produce a complaint or error, your installation has been successful!

### *Working with PyRosetta and PyMol on your local machine (for OS X and Linux users)*

After you have implemented the required methods in the MonteCarloPredictor in assignment 2, open PyMol. Using the PyMol GUI terminal, navigate to the assignment directory (using the `cd` command), and run the following:

```
PyMOL> run PyMOLPyRosettaServer.py
```

In a separate terminal window:

```
$ cd ~/Downloads/assn2 # navigate to your a2 dir
$ mkdir out # make a directory for your outputted pdb files to be stored
# If you've installed Pyrosetta using bash.rc/bash.profile (Method 2 in the previous section),
skip this
$ source ~/Downloads/PyRosetta-<version>/SetPyRosettaEnvironment.sh

# After running the command below, you should see the protein change conformation in Pymol
# Your pdb files should now be in your out directory
$ python predict.py pdbs/sequence.pdb dihedral1000.pdb -dihedral -pymol -1000
```

## Working on VPTL Macs

The VPTL Macs have all the necessary software packages already installed. To get started with PyRosetta, you must do this every time you log on to one of the VPTL machines:

1. Open up terminal and source the PyRosetta environment script as follows:

```
$ cd Downloads/assn2/  
$ source /Applications/PyRosetta/SetPyRosettaEnvironment.sh
```

You should get the following output:

```
Setting PyRosetta root as: /Applications/PyRosetta  
Aliasing PyRosetta Toolkit GUI to pyrosetta_toolkit  
/Profiles/<your suid>
```

You must do this **every time** you open up a new terminal window or tab.

### *Working with PyRosetta and PyMol on VPTL Machines*

Download and unzip the assn2.zip file from the class website and begin working on the assignment. After you have implemented the required methods in the MonteCarloPredictor methods in predictor.py, open up PyMol. Using the **PyMol GUI** terminal, navigate to the assignment directory (using the cd command), and run the following:

```
PyMOL> cd Downloads/assn2/  
cd: now in /afs/ir.stanford.edu/users/o/s/osamae/Desktop/assn2  
PyMOL> run PyMOLPyRosettaServer.py  
PyMOL <---> PyRosetta link started!  
at 127.0.0.1 port 65000
```

In a separate terminal window:

```
$ cd ~/Downloads/assn2 # navigate to your a2 dir  
$ mkdir out # make a directory for your outputted pdb files to be stored  
  
# After running the command below, you should see the protein change conformation in Pymol  
# Your pdb files should now be in your out directory  
$ python predict.py pdbs/sequence.pdb dihedral1000.pdb -dihedral -pymol -1000
```

## Working remotely (FarmVNC)

We have several rye machines you can work from remotely, but be warned that this workflow can be riddled with cryptic network problems.

First, you will need to use FarmVNC to set up an ssh tunnel to access the machines. Before beginning, follow the directions listed under [FarmVNC'S "First Steps"](#) to ensure that you have **AFS tokens** and a **VNC password** set up on Farmshare. You only have to set the VNC password once – be sure to remember it. Before beginning, you will also need a **VNC client**. Mac and Linux users should download and install [TurboVNC](#). Windows users should download and install [SS-VNC](#). Windows users will also need to download an ssh client like [PuTTY](#).

### *Establishing an SSH tunnel with FarmVNC*

**ssh** into one of the rye machines. Mac & Linux users can do this from the terminal. Windows users can do this using PuTTY:

```
$ ssh your-sunid@rye.stanford.edu
# PuTTY users would use rye.stanford.edu for the hostname
```

Next load the farmvnc module and run it. It will output a lot of initialization jargon – the important steps and information are highlighted.

```
rye01:~> module load farmvnc
rye01:~> farmvnc 1024x768 # choose appropriate resolution for your screen
checking for current vnc sessions on this host

checking port 5901
using display 1
searching for available GPU
GPU 0 available
using GPU 0

Initializing built-in extension Generic Event Extension
** Other initializations omitted **
Initializing built-in extension DRI2
Loading extension GLX
Loading extension NV-GLX
Loading extension NV-CONTROL

###
##
# Mac and Linux users
# run this command in another window to setup the SSH tunnel
#
# ssh -L 5901:localhost:5901 your-sunid@rye01.stanford.edu
#
# Windows users
# use these values in SS-VNC
#
# VNC Host Display: localhost:1
# Proxy/Gateway: your-sunid@rye01.stanford.edu
##
###
```

For Mac and Linux users, copy the instructed ssh [command](#), open up a new terminal window, and paste the command. After logging in, open up TurboVNC Viewer, and enter **localhost:5901**

into the VNC Server box. The next screen will ask you for your VNC password that you set earlier (different from your SUNET password).

For Windows users, open up **SS-VNC** and enter the highlighted **credentials**.

### *Using the VNC Client and Loading PyRosetta*

A window should have popped which gives you remote access to the rye machine. Open up the terminal. If you are just starting the assignment, you can copy the assignment files hosted on the website into your local space. The following will copy the zip file into a cs279/ directory in your home directory. You can use the unzip command to unzip compressed files.

```
rye01:~> cp /afs/ir/class/cs279/WWW/assignments/assn2.zip ~/cs279/  
rye01:~> cd cs279/  
rye01:~> unzip assn2.zip
```

To begin working on assignment 2, it is imperative that you load the pyrosetta module in the VNC Client's Terminal every time. PyMOL, on the other hand, can be launched by just typing PyMOL in the terminal.

```
rye01:~> module load pyrosetta
```

The last and most important step: **PLEASE log out of the VNC client by clicking on the gear icon at the top right corner of the virtual desktop and clicking log-out.** For good measure, kill all the processes under your SUNetID on the rye machine you using to ensure that you do not leave behind "zombie" processes:

```
rye03:~> pkill -u your-sunetid  
Connection to rye03.stanford.edu closed by remote host.  
Connection to rye03.stanford.edu closed.
```

You should be good to go for assignment 2!

### *Rye Machine Overloaded Troubleshooting*

It is possible, and at times, very likely that you will experience problems with the **rye** machines due to them being overloaded with "zombie" processes. Often times, these errors are cryptic and confusing. Here are the two most common errors you may encounter:

1. When you load the farmvnc module and execute the farmvnc command with specified screen resolution, you may get an error that the machine has no available GPUs.
2. When you launch the VNC client and enter the host credentials, you may get a weird abort message indicating the connection was refused (i.e. "channel 3: open failed: connect failed: Connection refused.")

If you get either of these errors, or are consistently running into odd connection refused error messages, this is a good indication that the rye machine you are using is overloaded with other students' processes. Your options here are somewhat limited:

1. Try using the other **rye** machines (rye01 through rye06). If you have no luck with one machine, be sure to kill the processes under your SUNetID before trying another (i.e. execute `pkill -u your-sunetid`)
2. Try installing and developing locally on your own machine
3. Wait a few hours. There is a script that runs in the middle of the night that cleans zombie processes

All that being said, it is imperative that you begin this assignment early to ensure that you can deal with any technical difficulties before the deadline!