

Installing Simulation Software

Jinko Graham

30/07/2019

- NOTE 1: This document includes code that you should cut-and-paste into your R console to install software. There is a second document called `testInst.Rmd` that you can use to test the installed software.
- NOTE 2: Before working through this installation document, you must copy, or “clone” the class GitHub repository to your computer. The simplest way to do this is to create an RStudio project linked to the git repository <https://github.com/SFUStatgen/SFUStat840>. See the instructions on the class Canvas page on Computing: <https://canvas.sfu.ca/courses/47316/pages/computing> for instructions on getting started with git and version control in RStudio.

Overview

We will work with simulation software called **msprime** and **SLiM**. **msprime** simulates ancestries of chromosomes according to a backward-in-time model called the coalescent that we will study in class. **SLiM** simulates genomic data forwards-in-time using a Wright-Fisher reproductive model that we will also study in class. Both programs return a data structure called a succinct tree sequence (STS) that represent ancestral relationships between chromosomes. **msprime** is a Python package (a.k.a. “module”) and **SLiM** is a C program. Installation of **msprime** also installs a toolkit called **tskit** that is useful manipulation of STSs in Python. A Python package called **pyslim** allows us to read STSs output by **SLiM** into Python so that we can manipulation them with the tools in **tskit**. The majority of this document is comprised of instructions for installing Python and the packages **msprime**, **tskit** and **pyslim**. At the end is a short section describing installation of **SLiM**.

Python and Python packages

I will be using an R package called **reticulate** to interface with Python and the Python packages. The following instructions will lead you through the installation of Python, **reticulate**, **msprime**, **tskit** and **pyslim**.

Install Python and conda

Conda is a package management system used by the Python community. Python and conda are installed simultaneously by a program called miniconda, which you can find at <https://docs.conda.io/en/latest/miniconda.html>

Choose the installer for Python version 3.7 that is appropriate for your operating system and follow the instructions to install. Note for Mac users: You can choose either a “bash installer” or a “.pkg installer”. I think the .pkg installer will be more familiar to most users. Download the .pkg file to your computer, double-click to start the installer and you will be guided through the installation.

- NB: The installation of conda changes the configuration of the shell used in your terminal windows. You need to open a *new* terminal after the installation to be able to use conda. From a new terminal, type **which python** and record the path to the python executable that was installed by miniconda (you will need it in the next step). On my Mac the path is `/Users/jgraham/miniconda3/bin/python`.

Install reticulate

This is an R package and can be installed from CRAN in the usual way. After installing, use `library()` to load the package and the function `use_python()` to set the path to your Python executable, as in the following R code chunk:

```
library(reticulate)
use_python("~/miniconda3/bin/python")
```

Installing msprime and tskit into a Python “environment”.

Python environments are like RStudio projects in that they are meant to encapsulate all code and data for a specific project. However, they take this encapsulation idea a step further and also include any Python package that you need for the project. In the following I create an environment called `Stat840` with the `conda_create()` function and install `msprime` and `tskit` into it with `conda_install()`. Since `msprime` depends on `tskit` (in much the same way as R packages can depend on each other), `tskit` will be automatically installed when we install `msprime`.

```
envname <- "Stat840"
conda_create(envname)
conda_install(envname, "msprime")
```

Installing pyslim

Unfortunately `pyslim` has not yet been made available on the `conda-forge` website that hosts conda-installable packages. I built a conda-installable version of `pyslim` myself and uploaded it to the class GitHub site. I’ve also written an function `conda_install_local()` to simplify the installation of `pyslim` from within R. Both the package and R function are in the class GitHub repository (<https://github.com/SFUStatgen/SFUStat840>). Before proceeding, you **must** have a copy of the repository on your computer. In what follows, replace `/path/to/SFUStat840` with the path to the copy of the class repository on your computer.

```
pkgname <- "pyslim"
dir <- "/path/to/SFUStat840/PythonPackages/"
source(paste0(dir, "conda_install_local.R"))
conda_install_local(envname, pkgname, dir)
```

Using the Stat840 Python environment in future R sessions

To use the `Stat840` Python environment and all the package you have installed into it, include the following code chunk at the top of your RMarkdown documents.

```
library(reticulate)
use_condaenv("Stat840")
```

Installing SLiM

SLiM is a C program available from <https://messerlab.org/slim/>

Unfortunately the program is only available for Mac OS X and Linux users. In fact, our group has only been successful installing SLiM on Mac OS X. If you are a Mac user, download and run the OS X installer from the above website. This should install both a command-line version of SLiM (e.g., in `/usr/local/bin/slim`) and a graphical-user-interface (GUI) version of SLiM (e.g., SLiMgui in your Applications folder).

If you do not have access to a Mac, you can get an account on a Mac Pro used by the Graham and McNeney labs. The computer is housed in Brad McNeney's office, and you will have to make an appointment with Brad to log in from the console the first time you use it. After the initial login, you can log in remotely. You will not be able to use the SLiM GUI, or any other GUI, when you log in remotely. Also, without GUI access you can't use RStudio, so you will need to use `git` from the command line to clone the repository to your account on this Mac; i.e., run `git clone https://github.com/SFUStatgen/SFUStat840`.