

MA2 + MA3 introduction

Multiprocessing, higher order functions, git, terminal

Andrey Shternshis¹

Department of Information Technology
Uppsala University

April 25, 2025

¹Courtesy of Sven-Erik Ekström

Data structures

You have learned how to create data structures using classes. You know how to initialize objects, iterate over elements, and write methods for insertion and removing.

Data structures are not limited by lists and trees (see hash tables), but now you know concepts and implementation techniques.

Queue

A queue is a collection of elements that are maintained in a sequence. It can be modified by the addition of elements at one end of the sequence and the removal of elements from the front.

The order is first in, first out (FIFO).

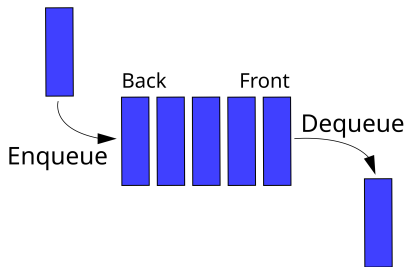


Figure: Queue. Vegpuff/Wikipedia, link

Stack

A stack is a collection of elements with two main operations:

- **Push** adds an element to the collection, and
- **Pop** removes the most recently added element.

The order is last in, first out (LIFO).

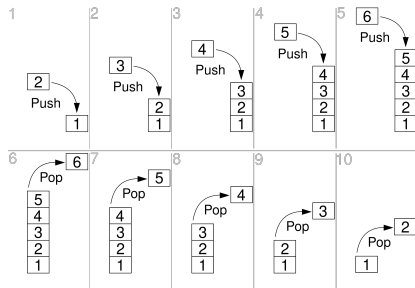


Figure: Stack, link

Efficient codes

Ways to run scripts faster

- ▶ Improve algorithm complexity
- ▶ Use multiprocessing: use two or more central processing units (CPU)
- ▶ Exploit Graphics processing units (GPU) accelerating computations
- ▶ Integrate a statically typed compiled language such as C++
- ▶ Rely on more python techniques and implementations: list comprehension and higher order functions

Benchmark problem

A problem to compare methods should have known solution to measure accuracy and be complicated enough to measure execution time.

- ▶ Estimation of π
- ▶ Checking that $2^{136279841} - 1$ is prime

Estimation of π

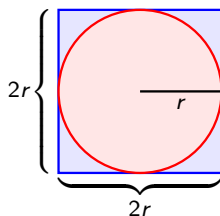


Figure: Circle with radius r inscribes in a square with sides $2r$.

A red circle with radius r and area $A_c = \pi r^2$. It is placed in a blue square, with sides $2r$ and that has area $A_s = (2r)^2 = 4r^2$.

Approximating π by Monte-Carlo methods

$$\frac{A_c}{A_s} = \frac{\pi r^2}{4r^2} = \frac{\pi}{4},$$

or

$$\pi = 4 \frac{A_c}{A_s}.$$

idea: Given that $r = 1$ and the center of the circle is at the origin, $(0,0)$, create n uniformly distributed random coordinates $(x,y) \in [-1,1] \times [-1,1]$ in the square.

$$\pi \approx 4 \frac{n_c}{n}.$$

where n_c points lie inside the circle

Alternative approaches

An alternative approach is in google colab using inverse squares.

The number of collisions of two cubes between themselves and a wall with energy consumption gives the first digits of π !

<https://youtu.be/HEfHFsfGXjs?si=U6uIpjN6gmRUtn00>

Volume of hyper-sphere

In 2D the volume is the area of a circle. In a dimension d , the volume of sphere is

$$V_d(r) = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} r^d$$

Question: What is $\Gamma(2)$?

To have accurate estimation of volume, a lot of samples are needed. It is time demanding especially in high dimensions.

Concepts and functionalities

List comprehension offers a short syntax when you want to create a new list based on the values of an existing list.

Higher order functions can take functions as parameters and return functions.

Multiprocessing

The speedup depends on your hardware

Parallel programming is not very efficient in Python due to counting for memory management

Different modules: 'multiprocessing' if nothing returned and 'concurrent.futures' if something returned

Linux machines

There are four Linux machines provided by IT department

- ▶ arrhenius.it.uu.se
- ▶ etc

To login write in the terminal:

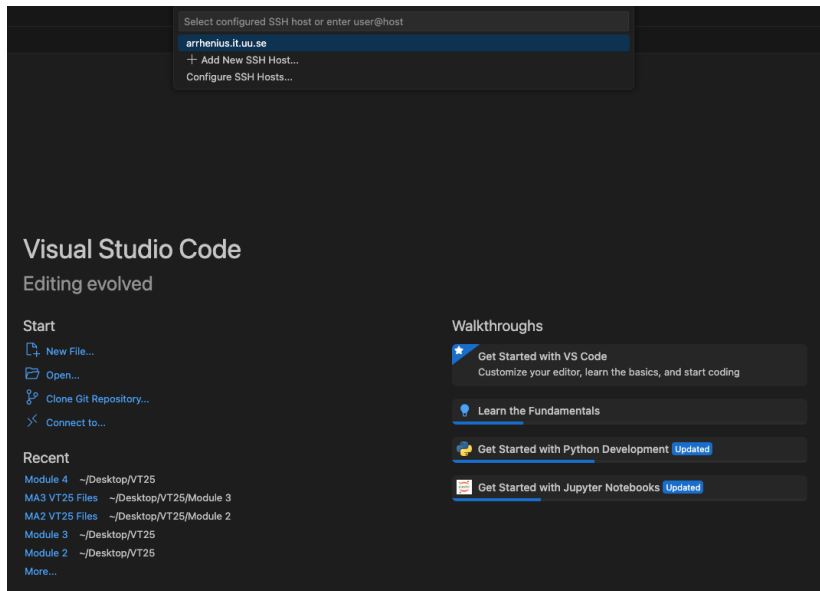
```
$ ssh abcde123@arrhenius.it.uu.se
```

\$ indicates that you have a terminal started. No need to write it.

The Secure Shell (SSH) protocol is a method for securely sending commands to a computer.

abcde123 is your login

Use graphical interface



Use linux commands

ls: List files in current directory (*ls -la* to show more information)

pwd: Show where you are (print working directory)

cd abc: Go into a directory named *abc*

cd ..: Go up one step in the file system (e.g., */* is the root directory, */home* contains home directories and */home/abcde123* is the home directory of user *abcde123*)

cd: Go to your home directory (where your personal files are, and this is where you are when you first log in)

mkdir abc: Create a directory called *abc*

rm -fr abc: Remove a file or directory called *abc*

nano hej.txt: Edit a file *hej.txt* with the editor *nano*

ctrl-x: Leave *nano*

python3 test.py: Run the Python code in the file *test.py* with Python 3.x.

Version control

Version control is the software engineering practice of controlling computer files and versions of files.

Reasons to use version control:

- ▶ Backups
- ▶ Code history
- ▶ Cooperation
- ▶ Work on several things (branches)
- ▶ Monitor when a code was broken
- ▶ Realise different versions to clients

Git

Git is a distributed version control system.

A repository is a directory where you keep code for a specific project.

One can either create repositories locally on your computer, or you can use providers of server hosted repositories.

- ▶ <https://github.com/>
- ▶ <https://gitlab.com/>
- ▶ <https://bitbucket.org/>

Access token

Repository can be open or private. You usually need private repository until the project is realized.

To clone from Github private repository to a server an access token should be provided due to security reasons.

Go to Settings → Developer Settings → Tokens → Generate new token.

Important to select access to commit status (tick repo:status)