MGE_downloader

by Christopher Schuster

This software is still in early testing. Do not run in production

MGE_downloader is tool that uploads FASTA files to the mobile genetic element finder and downloads the results. Uploads are staggered in order to prevent overloading of the server. Results are fetched after a certain wait time. The system uses Selenium, a browser control system. I chose this system as the uploads on the MEFinder website rely on Javascript queries and can easily be controlled with Selenium. All of this could probably be optimized. However, because only a few instances are run a the same time and the program is waiting most of the run-time, this would offer little advantage.

Please adhere to the minimum time between requests in order to no overload the server.

Installation of prerequesites

You will need to install Chrome, Chromedriver and Selenium first.

Windows Installation

- Install Chrome: https://www.google.com/chrome/
- Install Chromedriver:https://sites.google.com/a/chromium.org/chromedriver/downloads
- Install Python (3.7+): https://www.python.org/downloads/
- Install Selenium bindings through the command line: `pip install selenium`

Linux Installation

Ubuntu

- First, follow these instructions: https://gist.github.com/ziadoz/3e8ab7e944d02fe872c3454d17af31a5
- Then you will want to modify your ~/.bashrc to include the path to the script, e.g.

```
$ echo "PATH=$PATH:~/mge_downloader/" >> ~/.bashrc
$ source ~/.bashrc
```

ARCH

- Install chromium, python3 and selenium

```
$ sudo pacman -Syu
$ sudo pacman -S chromium python python-selenium
```

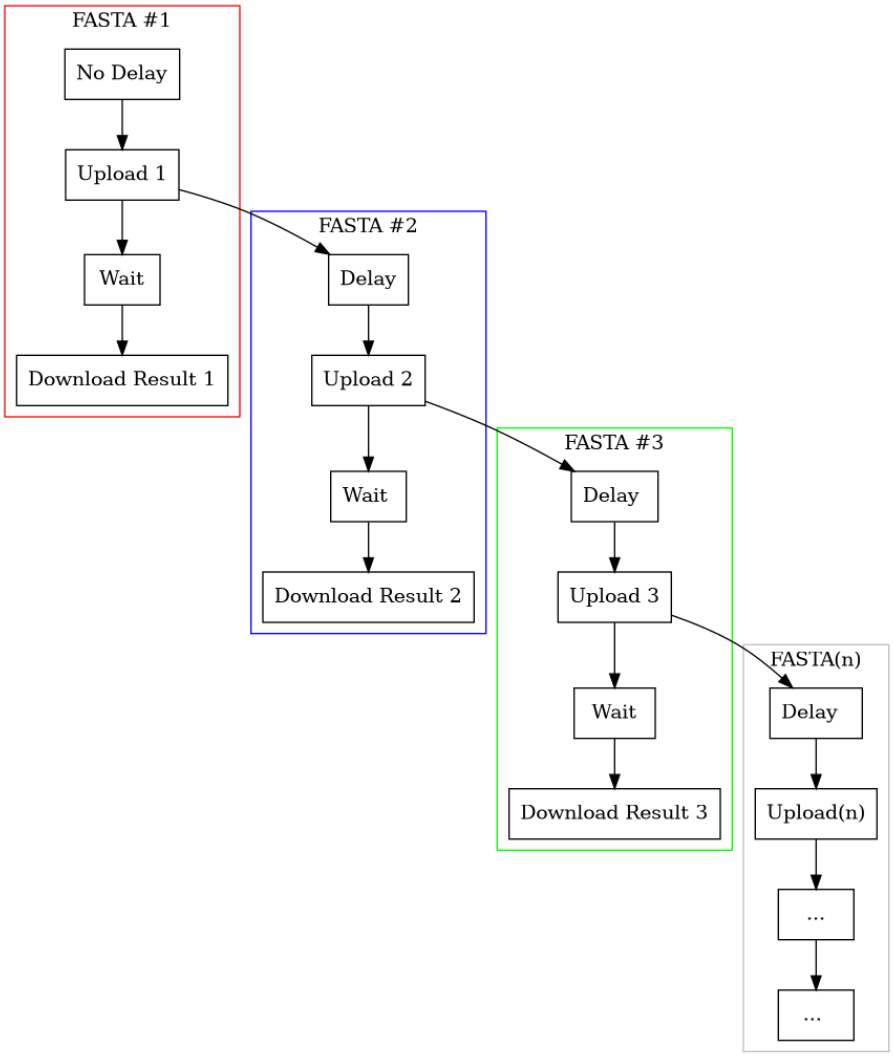
- Chromedriver can be installed through the AUR using `yay`

```
$ sudo pacman -S yay
$ yay -Syu
$ yay -S chromedriver
```

- Then you will want to modify your ~/.bashrc to include the path to the script, e.g.

```
$ echo "PATH=$PATH:~/mge_downloader/" >> ~/.bashrc
$ source ~/.bashrc
```

How it works



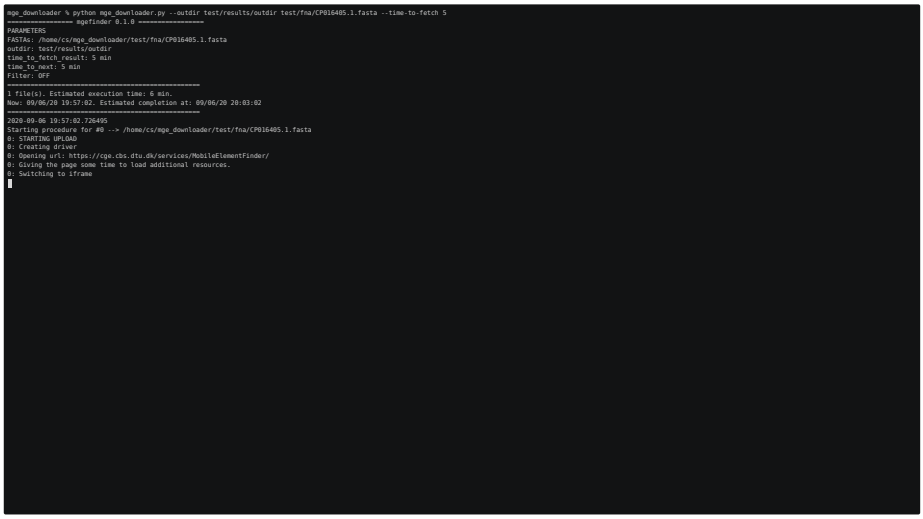
Workflow schema:

Quick Usage

Upload all fasta files from current directory and download results:

```
$ python mge_downloader --outdir outdir *.fna
```

Example usage



Usage

```
$ python main.py --help
================= mgefinder 0.1.0 =================
usage: main.py [-h] --outdir directory [--time-to-next minutes] [--time-to-fetch-result m

Upload fasta file to mefinder and download results.

positional arguments:
  fasta file(s)         fastA files.

optional arguments:
  -h, --help            show this help message and exit
  --outdir directory    directory that should be used to place the output.
  --time-to-next minutes
                        Number of minutes to wait before sending next fasta file
  --time-to-fetch-result minutes
                        Number of minutes to wait before fetching the result

Example: python mge_downloader.py --outdir dir1 *.fna
```

Contributing

Pull requests are welcome. For major changes, please open an issue first to discuss what you would like to change.

License

MIT