

CREOLE FORTH IN A JUPYTER NOTEBOOK



Joseph M. O'Connor
SVFIG MEETING
APRIL 2023



WHY DO IT

- Jupyter notebook is very commonly used in machine learning/data science.
- Very interactive and easy to use, like a one-dimensional spreadsheet.
- Supported in many different programming languages. Commonly Python, R, and Julia are used.
- Has lots of built-in tools.



IS FORTH SUPPORTED?

- Not directly, but Python is.
- That means a Forth written in Python can work.
- Fortunately, I've written a version for Python.
- We'll be taking a look at how it can be used with Jupyter notebook today.



INITIAL SETUP

- Import Creole Forth for Python along with two helper definitions to execute and compile the Forth code.
- Examples:
 - (1) `execCF('HELLO')` – executes the 'HELLO' primitive.
 - (2) `execCF('TEST')` – executes the 'TEST' primitive.
 - (3) `execCF('3 4 + .')` – adds two numbers, prints the result.
 - (4) `execCF('VLIST')` – lists the dictionary definitions.

LIMITATIONS TO THIS APPROACH

- Wrapping Forth code inside a Python function is cumbersome.
- It would be nice to do it more conveniently.



ONE ALTERNATIVE

- Create your own Jupyter notebook kernel.
- There are tools available such as Xeus which are designed for this.
- It's still a fair amount of work.



A SIMPLER SOLUTION

- Stick with the Python kernel.
- Python has a facility called magic commands, which allow the user to wrap a line or a cell inside a function and then call the function.
- It only requires writing a few lines of code.

THE MAGIC COMMANDS

- `%cfpy` – executes Forth commands on a single line.
- `%%cfpy` – executes Forth commands in a cell.
- `%compdef` – compiles Forth on a single line.
- `%%compdef` – compiles Forth in a cell.



SOME SIMPLE EXAMPLES

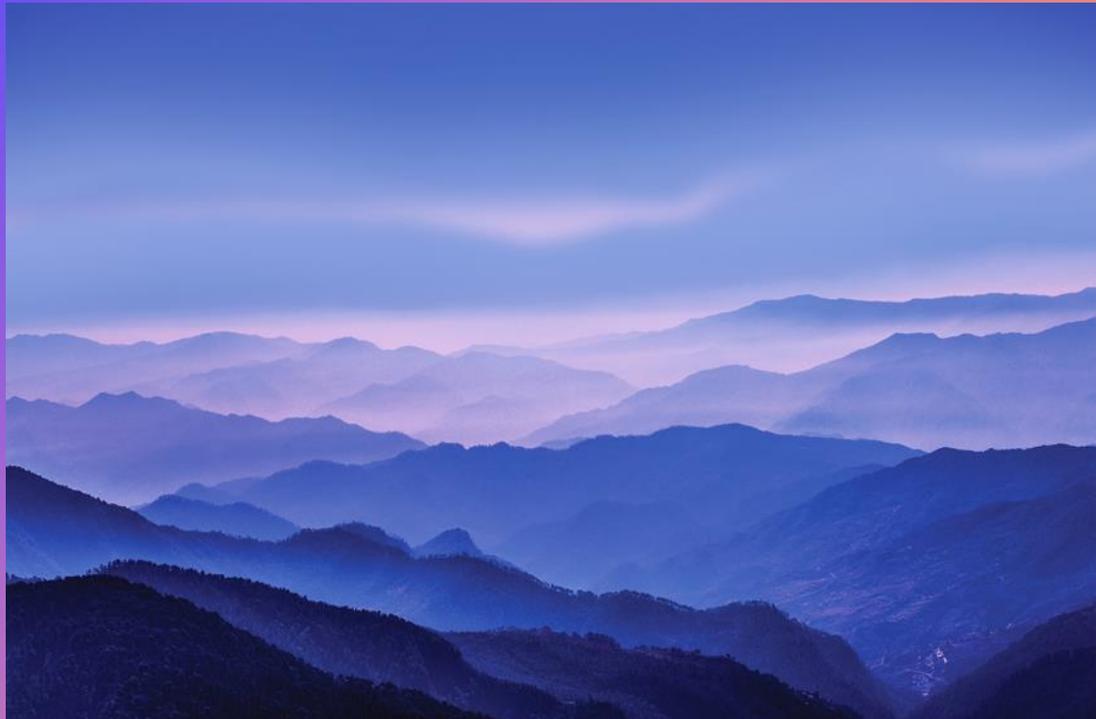
- `%cfpy HELLO` – Executes the HELLO primitive
- `%%cfpy` – Executes the code below the line.
- HELLO
- `%compdef : T2 TEST TEST ;`
- `%%compdef`
- `: TESTS 0 DO TEST LOOP ;`
- `%cfpy 3 TESTS`

EXAMPLE: MACHINE LEARNING

- Data analyzed is of thousands of urls which are classified as phishing, suspicious or legitimate.
- Exploratory data analysis, data cleaning, and looking for data correlations was initially done.
- It was followed up with binary classification to mark sites as phishing or non-phishing.
- First logistic regression was done with plots to show the effectiveness of the model.
- It was then validated with K-fold cross-validation.
- This is a methodology that resamples data in order to find the efficacy of machine learning models.
- Data is split into K subsamples. Each subsample is used as a testing set, while the remainder are used as training sets.
- It checks the performance of the model on new data in order to avoid overfitting or underfitting the model.

EXAMPLE – TODO LIST/LOG

- GUI front-end is built in Lazarus.
- It has two tabs, one for the list and the other for the log.
- Dialog box is called as an executable.
- The next cell executed has Creole Forth for Python code which uploads the saved text files to Dropbox.
- The files in Dropbox can then be viewed from any device with access to Dropbox, such as an iPad or Android.



Summary

1. Jupyter notebook is an effective IDE for interactive development.
2. A Forth written in Python can be adapted to use it without great effort.



QUESTIONS/COMMENTS?

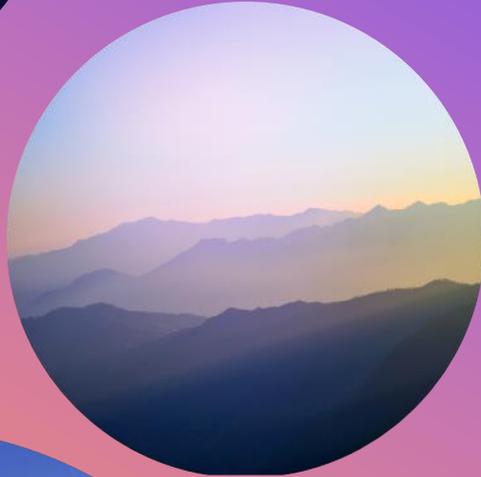
- Reach me at tiluser0@gmail.com
- Code for demo is available on Github at https://github.com/tiluser/cfpy_jn

PRESENTATION TITLE

+



o



.



THANK YOU

Joseph M. O'Connor
tiluser0@gmail.com