

Homework 8

Programming for Scientists

Due on: Apr 5

§1 QUESTIONS

Question I: What is a text file?

Question II: What does it mean to say that a file has “Unix line endings”?

Question III: What does the following regular expression match?

```
import re
pattern = re.compile(r'\d{2}-\d{2}-\d{2,4}')
```

This uses functionality not covered in class. Look up the `re` module.

§2 PROGRAMMING ASSIGNMENT

The following is the start of a file containing sparse arrays of numbers:

```
100 13
1:0.7 2:1 3:1 4:-0.3 7:-0.1
1:0.5 2:-1 3:0.333333 6:-0.603774 10:1
1:0.1 2:1 4:-0.333334 5:-0.433962 9:-0.383562 6:-1
1:0.4 2:1 4:1 4:-0.358491 10:0.374429 11:-1 12:-1
2:-1 3:-0.333333 4:-0.509434 5:-0.347032 6:-1 7:1
1:0.5 2:1 3:1 4:-0.509434 5:-0.767123 6:-1 7:-1 8:0.0534351
```

The exact format is as follows: the first line has two numbers: the number of arrays and the length of each array. Each subsequent line represents an array. The array is represented as a space-separated list of pairs *idx:value* where *idx* is the index on an entry and *value* is its value. Entries without a value are zero. For example, the first array is `(.7, 1, 1, -0.3, 0, 0, -1, 0, 0, 0, 0, 0)`

Write a function that parses this file and returns a numpy array with the values.