

Programming: The Next Step

Genome 559: Introduction to Statistical and
Computational Genomics

Elhanan Borenstein

A quick review

- Dictionaries:
 - key:value pairs
 - a.k.a. hash tables, lookup tables
 - Examples:
 - Word and definition
 - Name and phone number
 - Gene name and score
 - Username and password
 - Dictionaries are useful when you want to look up some data (value) based on a key
 - **Each key can appear only once**
- Standard I/O

Note: dictionary and list access times

- Accessing a list by index is very fast!
- Accessing a dictionary by key is very fast!
- Accessing a list by value (e.g. `list.index(myVal)` or `list.count(myVal)`) can be SLOW.



Take a deep breath ...

... and think how much you've learned!

4 weeks ago, this would have been gibberish:

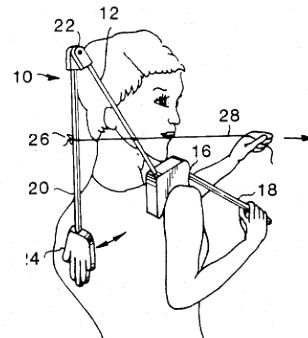
```
import sys
matrixFile = open(sys.argv[1], "r")
matrix = []
line = matrixFile.readline().strip()
while len(line) > 0:
    fields = line.split("\t")
    intList = []
    for field in fields:
        intList.append(int(field))
    matrix.append(intList)
    line = matrixFile.readline().strip()
matrixFile.close()

for row in matrix:
    for val in row:
        print val,
    print ""
```

initialize empty matrix
read first line stripped
until end of file
split line on tabs, giving a list of strings
create an int list to fill
for each field in current line
append the int value of field to intList
after intList is filled, append it to matrix
read next line and repeat loop

go through the matrix row by row
go through each value in the row
print each value without line break
add a line break after each row

FIG. 1



In theory,
what you know so far allows you
to solve any computational task
("universality")

So ... why don't we stop here?

**most real-life tasks
will be (very) painful to solve
using only what you know so far ...**

What are we missing?

What are we missing?

- A way to generalized procedures ...
- A way to store and handle complex data ...
- A way to organize our code ...
- Better design and coding practices ...

