Lecture 7

Python Standard Libraries

Homework

Just about the optimal answer (assuming the file isn't TOO big):

```
data=file(“infile.txt”,”r”).read()
data.replace(“It”,”IT”)
file(“outfile.txt”,”w”).write(data)
```

Homework

However, we really need to do this too:

```
data=file(“infile.txt”,”r”).read()
data.replace(“it”,”IT”)
data.replace(“It”,”IT”)
data.replace(“iT”,”IT”)
file(“outfile.txt”,”w”).write(data)
```

Unfortunately, this has the drawback of doing this:

“It is witty” ->
“IT is wITty”

Homework

We could try:

```
data=file(“infile.txt”,”r”).read()
data=data.split()
for i in range(len(data)) :
        if data[i].lower==”it” : data[i]=”IT”
file(“outfile.txt”,”w”).write(“ “.join(data))
```

This is pretty good, but:

```
“””It is witty,
don't you think it is ?”””  ->
“IT is witty, don't you think IT is ?”
```

Homework

How about this :

```
from re import *
data=file(“infile.txt”,”r”).read()
data=sub("(\A|\Z|(?<=\s))[iI][tT](\A|\Z|(?=\s))","IT",data)
file(“outfile.txt”,”w”).write(data)
```

Perfect, but now we need to talk about regular expressions...

Python Standard Libraries

Let's go over what you can do with a few of the more useful libraries:

- pprint
- string - constants
- os
- datetime
- time
- urllib

Interactive demo, see other file...

urllib

```
>>> from urllib import urlopen
>>> from pprint import pprint
>>> filein=urlopen(“http://docs.python.org/lib/module-urllib.html”)
>>> lines=filein.readlines()
>>> pprint(lines)
...
>>> lines=urlopen(“http://finance.yahoo.com/d/quotes.csv?
        s=GOOG&f=sl1d1t1c1ohgv&e=.csv”).readlines()
```

(no linebreak when you enter the above line)

```
>>> pprint(lines)
['"GOOG",421.66,"7/10/2006","10:32am",+1.21,423.59,425.23,421.50,1326767\r\n']
>>> pprint(lines[0].strip().split(','))
...
```

New Homework Assignment

- Due Friday (no class this thursday)
- Find something interesting on the internet, preferably something dynamic, and write a program to retrieve it, and if necessary, parse it. I will share the most interesting programs sent to me next lecture.