

# Ling 555 — Programming for Linguists

Python modules

Robert Albert Felty

Speech Research Laboratory  
Indiana University

Nov. 12, 2008

homework

modules

standard

## While we are waiting

```
svn checkout svn://robfelty.com\  
/home/robfelty/svn/l555 myl555
```

# Outline

homework

modules

standard

- 1 Homework questions and comments
- 2 More on modules
  - path
  - Tips
  - packages
- 3 Standard modules
  - System module
  - Os module
  - More user input — getopt
  - time
  - re

# The Python path

homework

In order to use modules, they must be in the path.

modules

python

```
import sys
sys.path
sys.path.append('mypydir')
```

path

Tips

packages

standard

# The Python path

homework

In order to use modules, they must be in the path.

modules

path

Tips

packages

standard

```
.bashrc
```

```
#prepend
```

```
export
```

```
    PYTHONPATH="mypydir:${PYTHONPATH}"
```

```
#append
```

```
export
```

```
    PYTHONPATH="${PYTHONPATH}:mypydir"
```

# Finding the source

homework

modules

path

Tips

packages

standard

To list the contents of a module, do:

```
import sys
dir(sys)
sys.__all__ # what would be imported
            using *
```

To find the source of a module, do:

```
import sys
sys.__file__
```

# Packages

homework

modules

path

Tips

packages

standard

## Definition

A package is a collection of modules, and has a special `__init__.py` file

# Argv — user input

homework

modules

standard

```
import sys  
print sys.argv
```

Look at argv.py for a few more examples

sys

os

getopt

time

re



# Os module

homework

modules

standard

sys

os

getopt

time

re

The os module provides an abstraction layer to make things work on different operating systems. For example, if you want to open a file in a subdirectory, the paths are different in UNIX and Windows.

```
import os
filename=os.path.join('texts',
                      'Dracula.txt')
```

# Getopt

homework

modules

standard

sys

os

getopt

time

re

The getopt module allows you to process options like most UNIX commands

See `opts.py`

homework

modules

standard

sys

os

getopt

time

re

## Example

- Write the time to a log file
- See how long it takes to do something (see `hmwk9.py`)

# Regular expressions

homework

modules

standard

sys

os

getopt

time

re

**search** perform a regex search

**match** perform a regex search from the beginning of a string

**sub** Substitute - default is globally, optionally you can specify count

**split** Split a string using regex

**compile** Pre-compile a regex (not necessary, but can boost speed)

**flags** Some flags you can pass, like ignore case