PROGRAMMING, DATA STRUCTURES AND ALGORITHMS IN PYTHON

Week 2, Lecture 2

Madhavan Mukund, Chennai Mathematical Institute
http://www.cmi.ac.in/~madhavan

Names, values and types

✳ Values have types

    ✳ Determine what operations are allowed

✳ Names inherit type from currently assigned value

    ✳ Can assign values of different types to a name

✳ `int, float, bool`

✳ `+,-,*,/,..   and,or,..   ==,!=,>,..`

Manipulating text

* Computation is a lot more than number crunching

* Text processing is increasingly important

  * Document preparation

  * Importing/exporting spreadsheet data

  * Matching search queries to content

Strings —type `str`

* Type string, `str`, a sequence of characters

  * A single character is a string of length 1

  * No separate type `char`

* Enclose in quotes—single, double, even triple!

```
city = 'Chennai'

title = "Hitchhiker's Guide to the Galaxy"

dialogue = '''He said his favourite book is
"Hitchhiker's Guide to the Galaxy"'''
```

Strings as sequences

* String: sequence or list of characters

* Positions 0,1,2,…,n-1 for a string of length n

  * `s = "hello"`

0	1	2	3	4
h	e	l	l	o

    -5 -4 -3 -2 -1

* Positions -1,-2,… count backwards from end

  * `s[1] == "e", s[-2] = "l"`

Operations on strings

* Combine two strings: concatenation, operator +

    * `s = "hello"`

    * `t = s + ", there"`

    * `t` is now `"hello, there"`

* `len(s)` returns length of `s`

* Will see other functions to manipulate strings later

Extracting substrings

A slice is a "segment" of a string

* `s = "hello"`

* `s[1:4]` is "ell"

* `s[i:j]` starts at `s[i]` and ends at `s[j-1]`

* `s[:j]` starts at `s[0]`, so `s[0:j]`

* `s[i:]` ends at `s[len(s)-1]`, so `s[i:len(s)]`

Modifying strings

* Cannot update a string "in place"

  * `s = "hello"`, want to change to `"help!"`

  * `s[3] = "p"` — error!

* Instead, use slices and concatenation

  * `s = s[0:3] + "p!"`

* Strings are immutable values (more later)

Summary

* Text values — type `str`, sequence of characters

  * Single character is string of length 1

* Extract individual characters by position

* Slices extract substrings

* `+` glues strings together

* Cannot update strings directly — immutable