

Objectives

- Wrap up lists, dictionaries
- Default parameters
- Writing documentation
- Broader Issue: Digital Humanities

Mar 19, 2010

Sprenkle - CSCI111

1

Review: Dictionaries

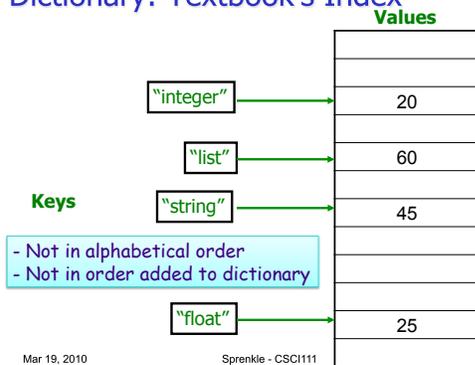
- What is the syntax for creating a new dictionary?
- How do we access a key's value from a dictionary?
 - What happens if there is no mapping for that key?
- How do we create a key → value mapping in a dictionary?

Mar 19, 2010

Sprenkle - CSCI111

2

Dictionary: Textbook's Index



Mar 19, 2010

Sprenkle - CSCI111

3

Creating Dictionaries in Python

Syntax:

```
{<key>:<value>, ..., <key>:<value>}
```

```
empty = {}
```

```
ascii = { 'a':97, 'b':98, 'c':99, ..., 'z':122 }
```

Mar 19, 2010

Sprenkle - CSCI111

4

Accessing Values Using Keys

- Typically, you should check if dictionary has a key before trying to access the key

```
if 'z' in ascii:  
    value = ascii['z']
```

Know mapping exists before trying to access

- Or handle if get default back

```
val = ascii.get('z')  
if val is None:  
    # do something ...
```

Mar 19, 2010

Sprenkle - CSCI111

5

Equivalent Solutions

```
if key not in dictionary :  
    dictionary[key] = 1  
else:  
    value = dictionary[key] + 1  
    dictionary[key] = value
```

```
if key not in dictionary :  
    dictionary[key] = 1  
else:  
    dictionary[key] += 1
```

Mar 19, 2010

Sprenkle - CSCI111

6

Defaults for Parameters

- Saw with the **xrange** function
 - Didn't have to specify start or increment when calling the function
 - Default start = 0
 - Default increment = 1
- Can assign a default value to a parameter
 - In general, in function header, default parameter (s) should come *after* all the parameters that need to be defined

Mar 19, 2010

Sprenkle - CSCI1111

7

Using Default Parameters

- By default, the rollDie function could assume that a die has 6 sides

Assigns a value to sides
ONLY IF not passed a parameter

```
def rollDie(sides=6):
    return random.randint(1,sides)
```

Examples of calling the function:

```
rollDie(6)
rollDie()
rollDie(12)
```

Show help

Mar 19, 2010

Sprenkle - CSCI1111

game.py

8

Problem: Student Majors

- We want to keep track of the number of majors of each type
 - Twist: Not every student has a major (don't declare until sophomore year)

Mar 19, 2010

Sprenkle - CSCI1111 majors_dictionary.py 9

Problem: Student Majors, revised

- Students can have more than one major
 - Should count these separately
- How can we modify the previous program to do that?

Mar 19, 2010

Sprenkle - CSCI1111

10

Lists vs. Dictionaries

Lists	Dictionaries
integer <i>positions</i> (0, ...) to any type of value	Map <i>immutable keys</i> (int, float, string) to any type of value
Ordered	Unordered
Slower to find a value (in)	Fast to find a value (use key)
Fast to print in order	Slower to print in order (by key)
Only as big as you make it	Takes up a lot of space (so can add elements in the middle)

Mar 19, 2010

Sprenkle - CSCI1111

11

Getting Documentation

- **dir**: function that returns a list of methods and attributes in an object
 - dir(<type>)
- **help**: get documentation
- In the Python shell
 - help(<type>)
 - import <modulename>
 - help(<modulename>)

Mar 19, 2010

Sprenkle - CSCI1111

12

Where is Documentation Coming From?

- Comes from the code itself in **“doc strings”**
 - i.e., “documentation strings”
- Doc strings are simply strings *after* the function header
 - Typically use triple-quoted strings because documentation goes across several lines

```
def verse(animal, sound):
    """ prints a verse of Old
    MacDonald, filling in the strings for
    animal and sound """
```

Mar 19, 2010

Sprengle - CSCI111

13

REPRESENTING DATA: NCAA BASKETBALL

Mar 19, 2010

Sprengle - CSCI111

14

March Madness, Baby!

- NCAA College Basketball Tournament
- 65 teams play for the championship
 - Play in several rounds
 - Duke - champ in 2001
- Broken into 4 brackets
 - 16 teams per bracket
 - Ranked 1-16
 - Games favor higher-ranked teams

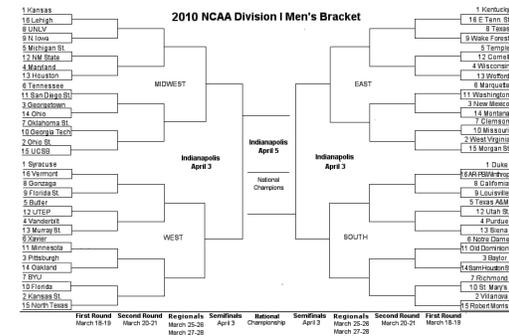


Mar 19, 2010

Sprengle - CSCI111

15

March Madness



Mar 19, 2010

Sprengle - CSCI111

16

Brainstorming on Tournament Info

- What data do we want to represent?
 - What data types would we use to represent this information?
- What do we want to be able to do with this data?

Mar 19, 2010

Sprengle - CSCI111

17

Challenge

- If you were allowed to pick 10,000 brackets, would one be a *perfect* bracket?
- What if you were given 100,000 brackets?

Mar 19, 2010

Sprengle - CSCI111

18

Digital Humanities

- Process large amounts of information to learn new information
 - Not necessarily numeric information
 - Text-based, images, multi-media, etc.
- Area growing in importance

Art:
James
Kelly Mae
George
Logan
Shannon

Art:
Nick
Dave
Sirocco
Dalena

Art:
Jeni
Collier
Taylor
Luke
Harrison

Metaphors:
Amy
Ben
Phil
Andrew
Will
Hank

Mar 19, 2010

Sprenkle - CSCI111

19

Discussion

- What problem were the researchers solving?
- How did they solve the problem?
 - What were some of the challenges in solving it?
- How “close” could they get to solving it?
 - When could their algorithms be wrong?
- Are there any follow up questions that can be asked/answered?
- What other gains do you see possible by applying computing to the humanities?
- Will digital humanities take the “art” out of the humanities?

Mar 19, 2010

Sprenkle - CSCI111

20

Relation to Class

- Analyzing texts, data
 - Strings
- Files containing the data
- Information retrieval, Natural Language Processing
 - Growing fields of computer science
 - E.g., Speech recognition
 - Process sentences
 - Determine subject/verb/direct object, etc.

Mar 19, 2010

Sprenkle - CSCI111

21

Learning Algorithms

- Train on data
 - This is an example of x
 - This is not an example of x
- Critical question: how to model/characterize info

Mar 19, 2010

Sprenkle - CSCI111

22

Analyzing Shakespeare

- hyphenated compound words
- Relative clauses per thousand
 - Shakespeare used less often
- Grade-level of writing, as measured by word- and sentence-length
- Percentage of open- and feminine-ended lines
 - Open: no punctuation
 - Feminine: unstressed syllable
 - Shakespeare used more often

Mar 19, 2010

Sprenkle - CSCI111

23

Next Week

- Monday: EC
 - Dr. Barry Trimmer of Tufts University
 - "Soft Bodies and Weak Minds: What Caterpillars Can Teach Us About Neuromechanics"
 - Stackhouse Theater (Commons) 10:10-11:05 a.m.

Mar 19, 2010

Sprenkle - CSCI111

24