University of Zurich[UZH]
Institute of Computational Linguistics

Programmiertechniken in der Computerlinguistik I

Herbstsemester 2015

6. Sitzung, 29. Oktober 2015

Martin Volk

University of Zurich[UZH]
Institute of Computational Linguistics

Python Tuples (are like Lists but immutable)

	Examples
Define a tuple - with three elements	my_tuple = () my_tuple = ('Junge', 'NN', 'ADJA')
Accessing an element	my_tuple[2]
Looping over the elements of a tuple	for x in my_tuple:
Tuples can be sorted	(2, 4, 100) < (2, 5, 1)

University of Zurich[UZH]

Institute of Computational Linguistics

Using tuples for sorting values from a hash

```
my_hash = {'x':3, 'z':10, 'y':4}
my_list = []
for key, value in my_hash.items():
        ## create a list of tuples
        my_list.append((value,key))

print sorted(my_list)
```

Page 3

University of Zurich[UZH]

Institute of Computational Linguistics

Python Functions

We define functions for program parts that we need to call repeatedly.

```
def my_print_function():
        for x in range(0, 30):
                print x, 'Hello world'

print 'And now we write some lines'
my_print_function()
print 'And again'
my_print_function()
```

Page 4

Python Functions

We can define and call functions with arguments.

```
def my_print_function(freq):
        for x in range(0, freq):
                print x, 'Hello world'

print 'And now we write some lines'
my_print_function(40)
```


Regular Expressions
Search and Substitution

	Examples
Regular expression matching	import re my_match = re.search('abc', my_text) my_match = re.search('^\waa..', my_text) print my_match.group()
Get hit from within a match	my_match = re.search('ab(\w+)', my_text) print my_match.group(1)
Get multiple matches in a list	my_matches = re.findall('a\w+', my_text)
Find and substitute	my_new_text = re.sub('t', 'th', my_text)

Programming Techniques in this Lesson

- Regular expression search and substitution
- Using Python tuples to sort hashed values
- Defining and calling Python functions
- Using a status variable to store a previous line
- Using a sentence-size window when iterating over a verticalized text
- Using two-dimensional lists for m*n matrices