# **PRG – PROGRAMMING ESSENTIALS**

Lecture 4 – Compound data types, Traversals <a href="https://cw.fel.cvut.cz/wiki/courses/be5b33prg/start">https://cw.fel.cvut.cz/wiki/courses/be5b33prg/start</a>

# **Tomas Jenicek**

Czech Technical University in Prague,
Faculty of Electrical Engineering, Dept. of Cybernetics,
Center for Machine Perception

http://cmp.felk.cvut.cz/~jenicto2/ tomas.jenicek@fel.cvut.cz

# **MEMORY**

```
x = 10
print(type(x))
                                                                                              11
y = x
if (id(x)==id(y)):
    print("x and y refer to the same object")
                                                                                              10
x = x + 1
if (id(x) != id(y)):
    print("x and y refer to DIFFERENT objects!")
z = 10
if (id(y)==id(z)):
    print("y and z point to the SAME memory!!")
else:
    print(" y and z point DIFFERENT objects!")
                                                                              Everything is object in Python
<class 'int'>
                                            Output Window
x and y refer to the same object
x and y refer to DIFFERENT objects!
y and z point to the SAME memory!!
```

# **MEMORY**



1 |

```
x = 10
print(type(x))
                                                                                               11
y = x
if (id(x)==id(y)):
    print("x and y refer to the same object")
x = x + 1
if (id(x) != id(y)):
    print("x and y refer to DIFFERENT objects!")
                                                                                               Car
z = 10
if (id(y)==id(z)):
    print("y and z point to the SAME memory!!")
else:
    print(" y and z point DIFFERENT objects!")
z = Car() #some user defined class
                                                                               Everything is object in Python
print(type(z))
<class 'int'>
                                            Output Window
                                                                              Python is a dynamically typed
x and y refer to the same object
                                                                                        language
x and y refer to DIFFERENT objects!
y and z point to the SAME memory!!
<class '__main__.Car'>
```

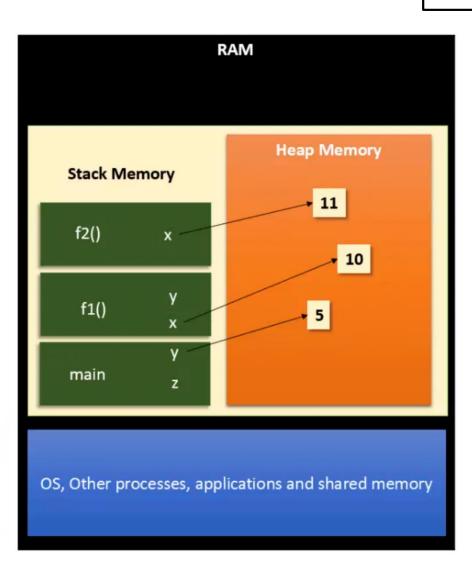
# **MEMORY**

```
#!/usr/bin/env.python
def.f1(x):
     x . *= . 2
     y_{x} = f2(x)
🗅 . . . return y
def . f2(x):
     x . += . 1
     return x
oif __name__ == '__main__':
     . y . = . 5
     z = 11(y)
     print(z)
 def f1(x):
```

```
def f1(x):

| Shadows name 'y' from outer scope less... (#F1)

This inspection detects shadowing names defined in outer scopes.
```





# **MORE ABOUT PYTHON**



	Python	JAVA / C
Statement	x = 10	int x = 10;
Data type declaration	Not needed. Dynamically typed.	Mandatory. Statically typed.
What is 10?	An Object created on heap memory.	A primitive data stored in 4 byte
What does x contain?	Reference to Object 10	Memory location where 10 is stored
x = x + 1	x starts referring to a new object whose value is 11	x continues to point to the same memory, with value equal to 11
x = 10 y = 10	Both x and y will refer to the same object.	x and y are two variables pointing to different memory locations.

# **MORE ABOUT PYTHON**

- Everything in Python is object
- Python is dynamically typed language (type changes with reference)
- The methods and variables are created on the stack memory
- The objects and instances are created on the heap memory
- New stack frame is created on invocation of a function / method and references are assigned & counted
- Stack frames are destroyed as soon as the function / method returns
- Mechanism to clean up the dead objects is Garbage collector (algorithm used is Reference Counting and immediate object removal if count == 0)





```
def mysum(xs):
 1
         """ Sum all the numbers in the list xs, and return the total.
         running total = 0
         for x in xs:
 5
             running total = running total + x
         return running total
     # Add tests like these to your test suite ...
     test(mysum([1, 2, 3, 4]) == 10)
 9
     test(mysum([1.25, 2.5, 1.75]) == 5.5)
10
     test(mysum([1, -2, 3]) == 2)
11
     test(mysum([ ]) == 0)
12
     test(mysum(range(11)) == 55) # 11 is not included in the list.
13
```

- Automate repetitive tasks without errors
- Repeated execution of a set of statements is called iteration
- Already explored for, now explore while
- Running through all items in a list is traversing / traversal



#### TRAVERSAL – THE WHILE LOOP



```
def sum_to(n):
    """ Return the sum of 1+2+3 ... n """
    ss = 0
    v = 1
    while v <= n:
        ss = ss + v
        v = v + 1
    return ss

# For your test suite
test(sum_to(4) == 10)
test(sum_to(1000) == 500500)</pre>
```

- The while statement has same meaning as in English
- Evaluate the condition (at line 5) either False or True.
- If the value is False, exit the while statement and continue execution at the next statement (line 8 in this case)
- If the value is True, execute each of the statements in the body (lines 6 and 7), then go back to the while statement



### TRAVERSAL – THE WHILE LOOP



n p

)

```
1
2
3
4
5
6
7
8
9
10
11
```

```
def sum_to(n):
    """ Return the sum of 1+2+3 ... n """
    ss = 0
    v = 1
    while v <= n:
        ss = ss + v
        v = v + 1
    return ss

# For your test suite
test(sum_to(4) == 10)
test(sum_to(1000) == 500500)</pre>
```

```
def sum_to(n):
    """ Return the sum of 1+2+3 ... n """
ss = 0
for v in range(n+1):
    ss = ss + v
return ss
```

- The while loop is more work than the equivalent for loop
- Need to manage the loop variable: give it an initial value, test for completion, update it in the body to enable termination
- Note: range generates a list up to but excluding the last value

# TRAVERSAL – WHILE vs. FOR

- Use a for loop if you know how many times the loop will execute (definite iteration — we know ahead some definite bounds for what is needed)
- Use a for to loop over iterables (to be explored in later classes) usually in combination with in
- Use while loop if you are required to repeat computation until given condition is met, and you cannot calculate in advance when this will happen (indefinite iteration — we do not know how many iterations will be needed)



### TRAVERSAL – BREAK vs. CONTINUE



```
while True:
1
       play the game once()
       response = input("Play again? (yes or no)")
        if response != "yes":
4
            break
    print("Goodbye!")
   for i in [12, 16, 17, 24, 29, 30]:
                                                     12
       if i % 2 == 1:  # If the number is odd
                                                     16
          continue # Don't process it
                                                     24
       print(i)
                                                     30
   print("done")
                                                     done
```

- The break statement in Python terminates the current loop and resumes execution at the next statement
- The continue statement in Python <u>returns the control to the</u> beginning of the current loop
- The continue statement <u>rejects all the remaining statements</u> in the current iteration of the loop ...

#### **EXAMPLE**



```
import random
                                  # We cover random numbers in the
 1
     rng = random.Random() # modules chapter, so peek ahead.
     number = rng.randrange(1, 1000) # Get random number between [1 and 1000).
 4
     guesses = 0
     msg = ""
     while True:
         guess = int(input(msg + "\nGuess my number between 1 and 1000: "))
         guesses += 1
10
11
         if guess > number:
12
             msg += str(guess) + " is too high.\n"
13
        elif guess < number:</pre>
             msg += str(guess) + " is too low.\n"
14
15
         else:
             break
16
17
```

- Guessing game
- This program makes use of the mathematical law of trichotomy (given real numbers a and b, exactly one of these three must be true: a > b, a < b, or a == b)</li>



### **COMPOUND DATA TYPES**



m p

- So far built-in types like int, float, bool
- <u>Compound data types</u>: strings, lists, dictionaries, and tuples are different from the others because they are made up of smaller pieces (characters in case of a string, items in case of a list)
- Types comprising smaller pieces are compound data types

#### PAIRED DATA



m p

```
print(celebs)
print(len(celebs)
3
[("Brad Pitt", 1963), ("Jack Nicholson", 1937), ("Justin Bieber", 1994)]
```

```
for (nm, yr) in celebs:
    if yr < 1980:
        print(nm)</pre>
```

```
Brad Pitt
Jack Nicholson
```

- Example of paired data: lists of names and lists of numbers
- Advanced way of representing data: making a pair of things is as simple as putting them into parentheses (i.e. tuples)

#### **NESTED DATA**



1 |

15

```
students = [
        ("John", ["CompSci", "Physics"]),
        ("Vusi", ["Maths", "CompSci", "Stats"]),
        ("Jess", ["CompSci", "Accounting", "Economics", "Management"]),
        ("Sarah", ["InfSys", "Accounting", "Economics", "CommLaw"]),
        ("Zuki", ["Sociology", "Economics", "Law", "Stats", "Music"])]
                                                            John takes 2 courses
   # Print all students with a count of their courses.
                                                            Vusi takes 3 courses
    for (name, subjects) in students:
                                                            Jess takes 4 courses
        print(name, "takes", len(subjects), "courses")
                                                            Sarah takes 4 courses
                                                            Zuki takes 5 courses
   # Count how many students are taking CompSci
   counter = 0
   for (name, subjects) in students:
       for s in subjects:
                                           # A nested Loop!
            if s == "CompSci":
                                        The number of students taking CompSci is 3
               counter += 1
6
    print("The number of students taking CompSci is", counter)
```

 Data structure — a mechanism for grouping and organizing data to make it easier to use



#### **TUPLES**



m p

```
>>> julia = ("Julia", "Roberts", 1967, "Duplicity", 2009, "Actress", "Atlanta, Georgia")
```

- The pair data example is an example of a tuple
- Tuple groups any number of items into a compound value
- Tuple is a comma-separated sequence of values
- Other languages often call it records
   (some related information that belongs together)
- <u>Important</u>: strings and tuples are <u>immutable</u> (once Python creates a tuple in memory, it cannot be changed)
- Elements of a tuple cannot be modified, new tuple holding different information should always be made instead!

#### **TUPLES**



```
(name, surname, b_year, movie, m_year, profession, b_place) = julia

>>> b = ("Bob", 19, "CS") # tuple packing

>>> (name, age, studies) = b # tuple unpacking
>>> name
'Bob'
>>> age

>>> (a, b, c, d) = (1, 2, 3)
ValueError: need more than 3 values to unpack
```

- Powerful tuple assignment (remember variable swapping?)
- Equivalent of multiple assignment statements
- Requirement: the number of variables on the left must match the number of elements in the tuple
- Tuple assignment is called tuple packing / unpacking

#### **TUPLES**

```
def f(r):
    """ Return (circumference, area) of a circle of radius r """
    c = 2 * math.pi * r
    a = math.pi * r * r
    return (c, a)
```

- Use of tuples in functions as return value
- Function can always only return a single value, but by making that value a tuple, as many values can be packed together as is needed (e.g. find the mean and the standard deviation)
- Tuple items can themselves be other tuples (nested tuples)
- Heterogeneous data structure: can be composed of elements of different types (tuples, strings, lists)

# **STRINGS**

```
if __name__ == '__main__':
    example = "Hello, world!"
      example.
```

- Example: upper is a method that can be invoked on any string object to create a new string, where all the characters are in uppercase
- lower, capitalize, swapcase ...
- Use documentation & help!

```
>>> fruit = "banana"
>>> m = fruit[1]
>>> print(m)
```

```
>>> m = fruit[0]
>>> print(m)
b
```

- Python uses square brackets to enclose the index indexing operator []
- The expression in brackets is called an index
- <u>Example</u>: The expression fruit[1] selects character number 1 from fruit, and creates a new string containing just this one character
- Computer scientists always start counting from zero!
- An index specifies a member of an ordered collection (in this case the collection of characters in the string)
- Index indicates which one you want, hence the name
- Index can be any integer expression (not only value)

# **INDEXING**



```
>>> fruit = "banana"
>>> list(enumerate(fruit))
[(0, 'b'), (1, 'a'), (2, 'n'), (3, 'a'), (4, 'n'), (5, 'a')]
```

```
>>> prime_nums = [2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31]
>>> prime_nums[4]
11
>>> friends = ["Joe", "Zoe", "Brad", "Angelina", "Zuki", "Thandi", "Paris"]
>>> friends[3]
'Angelina'
```

- Use enumerate to visualize indices
- Note that indexing strings returns a string: Python has no special type for a single character (string of length = 1)
- Use index to extract elements from a list



### **INDEXING**



m p

22

```
>>> fruit = "banana"
>>> len(fruit)
6
```

```
1  sz = len(fruit)
2  last = fruit[sz] # ERROR!
1  sz = len(fruit)
2  last = fruit[sz-1]
```

IndexError: string index out of range.

- Use len to extract the number of elements (indexing from 0!)
- Negative indices count backward from the end of the string
- The expression fruit[-1] yields the last letter
- Traversals: while vs. for comparison again!

```
ix = 0
while ix < len(fruit):
    letter = fruit[ix]
    print(letter)
    ix += 1</pre>
```

```
for c in fruit:
print(c)
```

### **SLICING**



```
>>> s = "Pirates of the Caribbean"
>>> print(s[0:7])
Pirates
>>> print(s[11:14])
the
>>> print(s[15:24])
Caribbean
>>> friends = ["Joe", "Zoe", "Brad", "Angelina", "Zuki", "Thandi", "Paris"]
>>> print(friends[2:4])
['Brad', 'Angelina']
```

- A substring of a string is obtained by taking a slice
- Slice a list to refer to some sublist of the items in the list
- The operator [n:m] returns the part of the string from the n'th character to the m'th character, including the first but excluding the last (indices pointing between the characters)
- Slice operator [n:m] copies out the part of the paper between the n and m positions
- Result of [n:m] will be of length (m-n)

### **SLICING**

```
>>> fruit = "banana"
>>> fruit[:3]
'ban'
>>> fruit[3:]
'ana'
>>> fruit[3:999]
'ana'
```

- If you omit the first index (before the colon), the slice <u>starts at</u> the beginning of the string (or list)
- If you omit the second index, the slice extends to the end of the string (or list)
- If you provide value for n that is bigger than the length of the string (or list), the slice will take all the values up to the end
- No "out of range" error like the normal indexing operation

### **STRINGS**



25

```
if word < "banana":
    print("Your word, " + word + ", comes before banana.")
elif word > "banana":
    print("Your word, " + word + ", comes after banana.")
else:
    print("Yes, we have no bananas!")
```

```
greeting = "Hello, world!"
greeting[0] = 'J' # ERROR!
print(greeting)

greeting = "Hello, world!"
new_greeting
print(reference)
```

greeting = "Hello, world!"
new\_greeting = "J" + greeting[1:]
print(new\_greeting)

- <u>Comparing strings</u>: strings are <u>sorted</u> in the alphabetical order (except that all uppercase letters come before the lowercase)
- Strings are immutable (existing string cannot be changed, new one should be created instead)

```
>>> "p" in "apple"
True
>>> "i" in "apple"
False
>>> "ap" in "apple"
True
>>> "pa" in "apple"
False
```

```
def find(strng, ch):
 1
 2
            Find and return the index of ch in strng.
            Return -1 if ch does not occur in strng.
          ix = 0
          while ix < len(strng):</pre>
               if strng[ix] == ch:
 9
                   return ix
               ix += 1
10
11
          return -1
12
     test(find("Compsci", "p") == 3)
13
     test(find("Compsci", "C") == 0)
test(find("Compsci", "i") == 6)
14
15
     test(find("Compsci", "x") == -1)
16
```

- The in / not in operator tests for membership
- Method index is the opposite of the indexing operator: it takes a character (item in case of a list) and finds the index of the character / item (if not found then exception is raised)
- Method find works for strings in a similar way (if the character is not found, the function returns -1)

# **STRINGS**



m

```
>>> ss = "Well I never did said Alice"
>>> wds = ss.split()
>>> wds
['Well', 'I', 'never', 'did', 'said', 'Alice']
```

- The split method:

   it splits a single multi-word string into a list of individual words, removing all the whitespace between them (whitespace are: tabs, newlines, spaces)
- Explore the join method on your own!

```
Python 3.6.9 (default, Sep 7 2019, 20:25:26)
In[2]: words = ['What', 'is', 'your', 'name', '?']
In[3]: sentence = ' '.join(words)

In[4]: print(sentence)
What is your name ?
In[5]:
```

### **STRINGS**



```
s1 = "His name is {0}!".format("Arthur")
     print(s1)
     name = "Alice"
                                                                   His name is Arthur!
     age = 10
     s2 = "I am {1} and I am {0} years old.".format(age, name)
                                                                   I am Alice and I am 10 years old.
     print(s2)
                                                                   2**10 = 1024 and 4 * 5 = 20.000000
9
    n1 = 4
10
     n2 = 5
     s3 = "2**10 = \{0\} \text{ and } \{1\} * \{2\} = \{3:f\}".format(2**10, n1, n2, n1 * n2)
12
     print(s3)
```

- The format method substitutes its arguments into the place holders (numbers are indexes of the arguments)
- Format specification it is always introduced by the colon:
- Field is aligned to the left <, center ^, or right >
- Width allocated to the field within the result string
- Type conversion
- Specification of decimal places (.2f is useful for when rounding to two decimal places.)

#### LISTS



n p

```
>>> vocabulary = ["apple", "cheese", "dog"]
>>> numbers = [17, 123]
>>> an_empty_list = []
>>> print(vocabulary, numbers, an_empty_list)
["apple", "cheese", "dog"] [17, 123] []
```

- A list is an ordered collection of values
- Values of a list are called its elements or items
- Similar to strings (ordered collections of characters)
   except that the elements of a list can be of any type
- Lists and strings and other collections that maintain the order of their items — are called sequences
- List within list is said to be nested
- List with no elements is called an empty list, and is denoted []

10

11

12

13

14

### **LISTS**



m p

```
students = [
    ("John", ["CompSci", "Physics"]),
    ("Vusi", ["Maths", "CompSci", "Stats"]),
    ("Jess", ["CompSci", "Accounting", "Economics", "Management"]),
    ("Sarah", ["InfSys", "Accounting", "Economics", "CommLaw"]),
    ("Zuki", ["Sociology", "Economics", "Law", "Stats", "Music"])]

# Count how many students are taking CompSci
counter = 0
for (name, subjects) in students:
    if "CompSci" in subjects:
        counter += 1

print("The number of students taking CompSci is", counter)
```

```
>>> a = [1, 2, 3]
>>> b = [4, 5, 6]
>>> c = a + b
>>> c
[1, 2, 3, 4, 5, 6]
```

```
>>> [0] * 4
[0, 0, 0, 0]
>>> [1, 2, 3] * 3
[1, 2, 3, 1, 2, 3, 1, 2, 3]
```

- Expression evaluating to an integer can be used as an index
- Function len returns length of a list (number of its elements)
- Testing membership using in / not in
- Operators + (concatenation) and \* (repetition)

#### **LISTS**



m p

```
>>> a_list = ["a", "b", "c", "d", "e", "f"]
>>> a_list[1:3]
['b', 'c']
>>> a_list[:4]
['a', 'b', 'c', 'd']
>>> a_list[3:]
['d', 'e', 'f']
>>> a_list[:]
['a', 'b', 'c', 'd', 'e', 'f']
```

```
>>> a_list = ["a", "d", "f"]
>>> a_list[1:1] = ["b", "c"]
>>> a_list
['a', 'b', 'c', 'd', 'f']
>>> a_list[4:4] = ["e"]
>>> a_list
['a', 'b', 'c', 'd', 'e', 'f']
```

```
>>> a_list = ["a", "b", "c", "d", "e", "f"]
>>> a_list[1:3] = []
>>> a_list
['a', 'd', 'e', 'f']
```

```
>>> my_string = "TEST"
>>> my_string[2] = "X"
Traceback (most recent call last):
   File "<interactive input>", line 1, in <module>
TypeError: 'str' object does not support item assignment
```

```
>>> my_list = ["T", "E", "S", "T"]
>>> my_list[2] = "X"
>>> my_list
['T', 'E', 'X', 'T']
```

- Lists are mutable (we can change list elements)
- Use same slicing principles as for strings
- Use del to delete list elements

### **REFERENCES – STRINGS vs. LISTS**



m p

- Variables a and b refer to string object with letters "banana"
- Use is operator or id function to find out the reference
- Strings are immutable
   Python optimizes resources by making two names that refer to the same string value refer to the same object
- Not the case of lists: a and b have the same value (content)
   but do not refer to the same object

# LISTS - ALIASING, CLONING



n p

```
>>> a = [1, 2, 3]

>>> b = a

>>> a is b

True

>>> b = [1, 2, 3]

>>> b = [1, 2, 3]

>>> b = a[:]

>>> b

[1, 2, 3]

>>> b = a[:]

>>> b

[1, 2, 3]

>>> a

[5, 2, 3]

>>> b = [1, 2, 3]

>>> a

[1, 2, 3]

>>> a

[1, 2, 3]
```

- If we assign one variable to another, both variables refer to the same object
- The same list has two different names we say that it is aliased (changes made with one alias affect the other)
- Recommendation:
   Avoid aliasing when you are working with mutable objects!
- If need to modify a list and keep a copy of the original use the slice operator (taking any slice of creates a new list)

# LIST PARAMETERS

m p

```
def double_stuff(a_list):
1
        """ Overwrite each element in a list with double its value.
       for (idx, val) in enumerate(a_list):
3
           a list[idx] = 2 * val
4
   things = [2, 5, 9]
                                     main
   double stuff(things)
                                                a list
   print(things)
                                                                  [2, 5, 9]
                              double_stuff
                                                things
 [4, 10, 18]
```

- Passing a list as an argument passes a reference to the list,
   not a copy or clone of the list!
- So parameter passing creates an alias!



### **LIST METHODS**



1 |

35

```
>>> mylist = []
>>> mylist.append(5)
>>> mylist.append(27)
>>> mylist.append(3)
>>> mylist.append(12)
>>> mylist
[5, 27, 3, 12]
```

```
>>> mylist.insert(1, 12) # Insert 12 at pos 1, shift other items up
>>> mylist
[5, 12, 27, 3, 12]
>>> mylist.count(12) # How many times is 12 in mylist?
>>> mylist.extend([5, 9, 5, 11]) # Put whole list onto end of mylist
>>> mylist
[5, 12, 27, 3, 12, 5, 9, 5, 11])
>>> mylist.index(9)
                                 # Find index of first 9 in mylist
6
>>> mylist.reverse()
>>> mylist
[11, 5, 9, 5, 12, 3, 27, 12, 5]
>>> mylist.sort()
>>> mylist
[3, 5, 5, 5, 9, 11, 12, 12, 27]
>>> mylist.remove(12)
                       # Remove the first 12 in the list
>>> mylist
[3, 5, 5, 5, 9, 11, 12, 27]
```

# **Explore list methods on your own!**



### LIST PARAMETERS



mρ

```
def double_stuff(a_list):
    """ Return a new list which contains
    doubles of the elements in a_list.

new_list = []
for value in a_list:
    new_elem = 2 * value
    new_list.append(new_elem)

return new_list
def double_stuff(a_list):
    """ Overwrite each element in a_list with double its value. """
for (idx, val) in enumerate(a_list):
    a_list[idx] = 2 * val
```

- Concept: pure functions vs. modifiers
- Pure function does not produce side effects!
- Pure function communicates with the calling program only through parameters (it does not modify) and a return value
- Do not alter the input parameters unless really necessary
- Programs that use pure functions are faster to develop and less error-prone than programs that use modifiers

### **REFERENCES**

m

37

#### This lecture re-uses selected parts of the OPEN BOOK PROJECT

**Learning with Python 3 (RLE)** 

http://openbookproject.net/thinkcs/python/english3e/index.html available under <u>GNU Free Documentation License</u> <u>Version 1.3</u>)

- Version date: October 2012
- by Peter Wentworth, Jeffrey Elkner, Allen B. Downey, and Chris Meyers (based on 2nd edition by Jeffrey Elkner, Allen B. Downey, and Chris Meyers)
- Source repository is at <a href="https://code.launchpad.net/~thinkcspy-rle-team/thinkcspy/thinkcspy3-rle">https://code.launchpad.net/~thinkcspy-rle-team/thinkcspy/thinkcspy3-rle</a>
- For offline use, download a zip file of the html or a pdf version from <a href="http://www.ict.ru.ac.za/Resources/cspw/thinkcspy3/">http://www.ict.ru.ac.za/Resources/cspw/thinkcspy3/</a>