Python 2.5 quick reference

John W. Shipman

2011-08-01 14:21

Abstract

A reference guide to most of the common features of the Python programming language, version 2.5.

This publication is available in Web form¹ and also as a PDF document². Please forward any comments to tcc-doc@nmt.edu.

Table of Contents

1. Introduction: What is Python? ................................................................. 5
2. Starting Python .......................................................................................... 5
   2.1. Using Python in Windows ................................................................. 5
   2.2. Using Python in Linux .................................................................... 5
3. Line syntax .................................................................................................. 6
4. Reserved words ............................................................................................ 6
5. Basic types .................................................................................................. 6
6. Numeric types ............................................................................................. 7
   6.1. Type int: Integers .............................................................................. 7
   6.2. Type long: Extended-precision integers .......................................... 8
   6.3. Type bool: Boolean truth values ..................................................... 8
   6.4. Type float: Floating-point numbers ............................................... 9
   6.5. Type complex: Imaginary numbers ............................................... 9
7. Sequence types .......................................................................................... 10
   7.1. Operations common to all the sequence types ................................ 10
   7.2. Type str: Strings of 8-bit characters .............................................. 13
      7.2.1. String constants ........................................................................ 13
      7.2.2. The string format operator ...................................................... 14
      7.2.3. String formatting from a dictionary ...................................... 16
      7.2.4. Definition of “whitespace” .................................................... 16
      7.2.5. Methods on str values ........................................................... 17
   7.3. Type unicode: Strings of 32-bit characters ..................................... 24
      7.3.1. The UTF-8 encoding .............................................................. 24
   7.4. Type list: Mutable sequences ......................................................... 25
      7.4.1. Methods on lists ..................................................................... 26
      7.4.2. List comprehensions ............................................................. 29
   7.5. Type tuple: Immutable sequences ................................................. 30
8. Types set and frozenset: Set types ....................................................... 31
   8.1. Operations on mutable and immutable sets .................................... 32

¹ http://www.nmt.edu/tcc/help/pubs/python25/web/
<table>
<thead>
<tr>
<th>Number</th>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1</td>
<td>abs()</td>
<td>Absolute value</td>
</tr>
<tr>
<td>13.2</td>
<td>all()</td>
<td>Are all the elements of a sequence true?</td>
</tr>
<tr>
<td>13.3</td>
<td>any()</td>
<td>Are any of the members of a sequence true?</td>
</tr>
<tr>
<td>13.4</td>
<td>bool()</td>
<td>Convert to Boolean</td>
</tr>
<tr>
<td>13.5</td>
<td>chr()</td>
<td>Get the character with a given code</td>
</tr>
<tr>
<td>13.6</td>
<td>cmp()</td>
<td>Compare two values</td>
</tr>
<tr>
<td>13.7</td>
<td>complex()</td>
<td>Convert to complex type</td>
</tr>
<tr>
<td>13.8</td>
<td>dict()</td>
<td>Convert to a dictionary</td>
</tr>
<tr>
<td>13.9</td>
<td>divmod()</td>
<td>Quotient and remainder</td>
</tr>
<tr>
<td>13.10</td>
<td>enumerate()</td>
<td>Step through indices and values of a sequence</td>
</tr>
<tr>
<td>13.11</td>
<td>file()</td>
<td>Open a file</td>
</tr>
<tr>
<td>13.12</td>
<td>filter()</td>
<td>Extract qualifying elements from a sequence</td>
</tr>
<tr>
<td>13.13</td>
<td>float()</td>
<td>Convert to float type</td>
</tr>
<tr>
<td>13.14</td>
<td>frozenset()</td>
<td>Create a frozen set</td>
</tr>
<tr>
<td>13.15</td>
<td>hex()</td>
<td>Convert to base 16</td>
</tr>
<tr>
<td>13.16</td>
<td>int()</td>
<td>Convert to int type</td>
</tr>
<tr>
<td>13.17</td>
<td>iter()</td>
<td>Produce an iterator over a sequence</td>
</tr>
<tr>
<td>13.18</td>
<td>len()</td>
<td>Number of elements</td>
</tr>
<tr>
<td>13.19</td>
<td>list()</td>
<td>Convert to a list</td>
</tr>
<tr>
<td>13.20</td>
<td>long()</td>
<td>Convert to long type</td>
</tr>
<tr>
<td>13.21</td>
<td>map()</td>
<td>Apply a function to each element of a sequence</td>
</tr>
<tr>
<td>13.22</td>
<td>max()</td>
<td>Largest element of a sequence</td>
</tr>
<tr>
<td>13.23</td>
<td>min()</td>
<td>Smallest element of a sequence</td>
</tr>
<tr>
<td>13.24</td>
<td>oct()</td>
<td>Convert to base 8</td>
</tr>
<tr>
<td>13.25</td>
<td>open()</td>
<td>Open a file</td>
</tr>
<tr>
<td>13.26</td>
<td>ord()</td>
<td>Find the numeric code for a character</td>
</tr>
<tr>
<td>13.27</td>
<td>pow()</td>
<td>Exponentiation</td>
</tr>
<tr>
<td>13.28</td>
<td>range()</td>
<td>Generate an arithmetic progression as a list</td>
</tr>
<tr>
<td>13.29</td>
<td>raw_input()</td>
<td>Prompt and read a string from the user</td>
</tr>
<tr>
<td>13.30</td>
<td>reduce()</td>
<td>Sequence reduction</td>
</tr>
<tr>
<td>13.31</td>
<td>reversed()</td>
<td>Produce a reverse iterator</td>
</tr>
<tr>
<td>13.32</td>
<td>round()</td>
<td>Round to the nearest integral value</td>
</tr>
<tr>
<td>13.33</td>
<td>set()</td>
<td>Create an algebraic set</td>
</tr>
<tr>
<td>13.34</td>
<td>sorted()</td>
<td>Sort a sequence</td>
</tr>
<tr>
<td>13.35</td>
<td>str()</td>
<td>Convert to str type</td>
</tr>
<tr>
<td>13.36</td>
<td>sum()</td>
<td>Total the elements of a sequence</td>
</tr>
<tr>
<td>13.37</td>
<td>tuple()</td>
<td>Convert to a tuple</td>
</tr>
<tr>
<td>13.38</td>
<td>type()</td>
<td>Return a value's type</td>
</tr>
<tr>
<td>13.39</td>
<td>unichr()</td>
<td>Convert a numeric code to a Unicode character</td>
</tr>
<tr>
<td>13.40</td>
<td>unicode()</td>
<td>Convert to a Unicode string</td>
</tr>
<tr>
<td>13.41</td>
<td>xrange()</td>
<td>Arithmetic progression generator</td>
</tr>
<tr>
<td>13.42</td>
<td>zip()</td>
<td>Combine multiple sequences</td>
</tr>
</tbody>
</table>

8.2. Operations on mutable sets ................................................................. 34
9. Type dict: Dictionaries ............................................................................. 36
9.1. Operations on dictionaries .................................................................. 37
10. Type file: Input and output files .......................................................... 40
10.1. Methods on file objects ...................................................................... 42
11. None: The special placeholder value .................................................... 43
12. Operators and expressions ..................................................................... 44
12.1. What is a predicate? ............................................................................ 44
13. Basic functions .......................................................................................... 45
13.1. abs(): Absolute value .......................................................................... 45
13.2. all(): Are all the elements of a sequence true? ................................... 45
13.3. any(): Are any of the members of a sequence true? ............................. 45
13.4. bool(): Convert to Boolean .................................................................. 46
13.5. chr(): Get the character with a given code ......................................... 46
13.6. cmp(): Compare two values ................................................................... 46
13.7. complex(): Convert to complex type .................................................... 47
13.8. dict(): Convert to a dictionary ........................................................... 47
13.9. divmod(): Quotient and remainder ..................................................... 48
13.10. enumerate(): Step through indices and values of a sequence .............. 48
13.11. file(): Open a file ............................................................................ 48
13.12. filter(): Extract qualifying elements from a sequence ...................... 48
13.13. float(): Convert to float type ............................................................ 49
13.14. frozenset(): Create a frozen set ......................................................... 49
13.15. hex(): Convert to base 16 .................................................................. 49
13.16. int(): Convert to int type ................................................................... 50
13.17. iter(): Produce an iterator over a sequence ....................................... 50
13.18. len(): Number of elements .................................................................. 50
13.19. list(): Convert to a list ...................................................................... 51
13.20. long(): Convert to long type ............................................................... 51
13.21. map(): Apply a function to each element of a sequence ...................... 51
13.22. max(): Largest element of a sequence ................................................ 51
13.23. min(): Smallest element of a sequence ................................................ 52
13.24. oct(): Convert to base 8 ..................................................................... 52
13.25. open(): Open a file ............................................................................ 52
13.26. ord(): Find the numeric code for a character ....................................... 52
13.27. pow(): Exponentiation ......................................................................... 52
13.28. range(): Generate an arithmetic progression as a list ....................... 53
13.29. raw_input(): Prompt and read a string from the user ........................... 53
13.30. reduce(): Sequence reduction ............................................................. 54
13.31. reversed(): Produce a reverse iterator ................................................ 54
13.32. round(): Round to the nearest integral value ....................................... 55
13.33. set(): Create an algebraic set .............................................................. 55
13.34. sorted(): Sort a sequence .................................................................... 56
13.35. str(): Convert to str type .................................................................... 56
13.36. sum(): Total the elements of a sequence ............................................. 57
13.37. tuple(): Convert to a tuple ................................................................. 57
13.38. type(): Return a value’s type ............................................................... 57
13.39. unichr(): Convert a numeric code to a Unicode character ................. 58
13.40. unicode(): Convert to a Unicode string ............................................... 58
13.41. xrange(): Arithmetic progression generator ....................................... 58
13.42. zip(): Combine multiple sequences ..................................................... 59

14. Advanced functions ...................................................................................... 59
14.1. basestring: The string base class ................................................................. 59
14.2. callable(): Is this thing callable? ................................................................. 59
14.3. classmethod(): Create a class method ...................................................... 60
14.4. delattr(): Delete a named attribute .......................................................... 60
14.5. dir(): Display a namespace's names ......................................................... 60
14.6. eval(): Evaluate an expression in source form ........................................... 61
14.7. execfile(): Execute a Python source file .................................................. 62
14.8. getattr(): Retrieve an attribute of a given name ...................................... 62
14.9. globals(): Dictionary of global name bindings ......................................... 62
14.10. hasattr(): Does a value have an attribute of a given name? .................... 63
14.11. id(): Unique identifier ............................................................................. 63
14.12. isinstance(): Is a value an instance of some class or type? ..................... 63
14.13. issubclass(): Is a class a subclass of some other class? ......................... 64
14.14. locals(): Dictionary of local name bindings ............................................. 64
14.15. property(): Create an access-controlled attribute .................................... 65
14.16. reload(): Reload a module ...................................................................... 66
14.17. repr(): Representation ............................................................................ 66
14.18. setattr(): Set an attribute ....................................................................... 67
14.19. slice(): Create a slice instance ................................................................. 67
14.20. staticmethod(): Create a static method .................................................. 68
14.21. super(): Superclass .................................................................................. 68
14.22. vars(): Local variables ............................................................................. 68

15. Simple statements .......................................................................................... 68
15.1. The assignment statement: name = expression ......................................... 69
15.2. The assert statement: Verify preconditions .............................................. 72
15.3. The del statement: Delete a name or part of a value .................................. 73
15.4. The exec statement: Execute Python source code .................................... 73
15.5. The global statement: Declare access to a global name ......................... 73
15.6. The import statement: Use a module ......................................................... 75
15.7. The pass statement: Do nothing .................................................................. 76
15.8. The print statement: Display output values .............................................. 76

16. Compound statements ................................................................................... 76
16.1. Python's block structure ............................................................................ 77
16.2. The break statement: Exit a for or while loop ....................................... 78
16.3. The continue statement: Jump to the next cycle of a for or while .......... 78
16.4. The for statement: Iteration over a sequence ......................................... 79
16.5. The if statement: Conditional execution ................................................... 80
16.6. The raise statement: Cause an exception .............................................. 81
16.7. The return statement: Exit a function or method .................................... 82
16.8. The try statement: Anticipate exceptions .............................................. 82
16.9. The yield statement: Generate one result from a generator ................. 85

17. def(): Defining your own functions ............................................................. 85
17.1. A function's local namespace ................................................................. 87
17.2. Iterators: Values that can produce a sequence of values ......................... 88
17.3. Generators: Functions that can produce a sequence of values ............... 89

18. Exceptions: Error signaling and handling ..................................................... 90
18.1. Definitions of exception terms ................................................................... 90
18.2. Life cycle of an exception ....................................................................... 91
18.3. Built-in exceptions ................................................................................... 91

19. Classes: Defining your own types ................................................................. 93
19.1. Old-style classes ....................................................................................... 96
19.1.1. Defining an old-style class ................................................................. 97
19.1.2. Instantiation of an old-style class: The constructor, __init__() .................................................. 97
19.1.3. Attribute references in old-style classes ................................................................. 97
19.1.4. Method calls in an old-style class ........................................................................ 98
19.1.5. Instance deletion: the destructor, __del__() ........................................................ 99
19.2. Life cycle of a new-style class ...................................................................................... 100
19.2.1. __new__(): New instance creation .................................................................... 100
19.2.2. Attribute access control in new-style classes ...................................................... 100
19.2.3. Properties in new-style classes: Fine-grained attribute access control ......... 101
19.2.4. Conserving memory with __slots__ .................................................................. 101
19.3. Special method names ............................................................................................... 102
19.3.1. Rich comparison methods .................................................................................. 103
19.3.2. Special methods for binary operators ................................................................ 103
19.3.3. Unary operator special methods ......................................................................... 104
19.3.4. Special methods to emulate built-in functions .................................................... 104
19.3.5. __call__(): What to do when someone calls an instance .................................. 105
19.3.6. __cmp__(): Generalized comparison .................................................................. 105
19.3.7. __contains__(): The “in” and “not in” operators ........................................... 106
19.3.8. __del__(): Destructor ..................................................................................... 106
19.3.9. __delattr__(): Delete an attribute .................................................................... 106
19.3.10. __delitem__(): Delete one item of a sequence ............................................. 106
19.3.11. __getattr__(): Handle a reference to an unknown attribute ...................... 106
19.3.12. __getattribute__(): Intercept all attribute references ................................ 107
19.3.13. __getitem__(): Get one item from a sequence or mapping ..................... 107
19.3.14. __iter__(): Create an iterator .......................................................................... 107
19.3.15. __nonzero__(): True/false evaluation .............................................................. 107
19.3.16. __repr__(): String representation .................................................................... 107
19.3.17. __setattr__(): Intercept all attribute changes ........................................... 108
19.3.18. __setitem__(): Assign a value to one item of a sequence .......................... 108
19.4. Static methods ........................................................................................................... 108
19.5. Class methods .......................................................................................................... 109
20. pdb: The Python interactive debugger .......................................................................... 109
20.1. Starting up pdb ........................................................................................................ 109
20.2. Functions exported by pdb ...................................................................................... 110
20.3. Commands available in pdb .................................................................................. 110
21. Commonly used modules .............................................................................................. 112
21.1. math: Common mathematical operations .................................................................. 112
21.2. string: Utility functions for strings ......................................................................... 113
21.3. random: Random number generation ...................................................................... 114
21.4. time: Clock and calendar functions ......................................................................... 115
21.5. re: Regular expression pattern-matching .................................................................. 117
21.5.1. Characters in regular expressions ...................................................................... 117
21.5.2. Functions in the re module ............................................................................. 118
21.5.3. Compiled regular expression objects .................................................................. 119
21.5.4. Methods on a MatchObject .............................................................................. 120
21.6. sys: Universal system interface ............................................................................... 120
21.7. os: The operating system interface ......................................................................... 121
21.8. stat: Interpretation of file status ............................................................................ 123
21.9. os.path: File and directory interface ...................................................................... 125
1. Introduction: What is Python?

Python is a recent, general-purpose, high-level programming language. It is freely available and runs pretty much everywhere.

- This document is a reference guide, not a tutorial. If you are new to Python programming, see A Python programming tutorial.3
- Complete documentation and free installs are available from the python.org homepage4.

This document does not describe every single feature of Python 2.5. A few interesting features that 99% of Python users will never need, such as metaclasses, are not described here. Refer to the official documentation for the full feature set.

2. Starting Python

You can use Python in two different ways:

- In “calculator” or “conversational mode”, Python will prompt you for input with three greater-than signs (>>>). Type a line and Python will print the result. Here's an example:

  ```
  >>> 2+2
  4
  >>> 1.0 / 7.0
  0.14285714285714285
  ```

- You can also use Python to write a program, sometimes called a script.

2.1. Using Python in Windows

If you are using Python at the NM Tech Computer Center (TCC), you can get conversational mode from Start → All Programs → ActiveState ActivePython 2.5 → Python Interactive Shell. To write a program:

1. Start → All Programs → ActiveState ActivePython 2.5 → PythonWin Editor.
2. Use File → New, select Python Script in the pop-up menu, and click OK. This will bring up an edit window.
3. Write your Python program in the edit window, then use File → Save As... to save it under some file name that ends in “.py”.
4. To run your program, use File → Run. In the “Run Script” popup, enter the name of your program in the field labeled Script File, then click OK.

The output will appear in the “Interactive Window”.

2.2. Using Python in Linux

To enter conversational mode on a Linux system, type this command:

```
python
```

---

4 http://www.python.org/
Type *Control-D* to terminate the session.

If you write a Python script named `filename.py`, you can execute it using the command

```
python filename.py
```

Under Unix, you can also make a script self-executing by placing this line at the top:

```
#!/usr/local/bin/python
```

You must also tell Linux that the file is executable by using the command “chmod +x filename”. For example, if your script is called `hello.py`, you would type this command:

```
chmod +x hello.py
```

### 3. Line syntax

The comment character is “#”; comments are terminated by end of line.

Long lines may be continued by ending the line with a backslash (\), but this is not necessary if there is at least one open “(”, “[”, or “{”.

### 4. Reserved words

```
and, assert, break, class, continue, def, del, elif, else, except, exec, finally, for, from, global, if, import, in, is, lambda, not, or, pass, print, raise, return, try, while, and yield
```

### 5. Basic types

In programming, you manipulate *values* using *operators*. For example, in the expression “1+2”, the addition operator (+) is operating on the values 1 and 2 to produce the sum, 3. The Python operators are described in Section 12, “Operators and expressions” (p. 44), but let’s look first at Python’s way of representing values.

Every Python value must have a type. For example, the type of the whole number 1 is `int`, short for “integer.”

Here is a table summarizing most of the commonly-used Python types.
Table 1. Python's common types

<table>
<thead>
<tr>
<th>Type name</th>
<th>Values</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>int</td>
<td>Integers in the range ([-2^{31}, 2^{31}-1]). See Section 6.1, “Type int: Integers” (p. 7).</td>
<td>42, -3, 1000000</td>
</tr>
<tr>
<td>long</td>
<td>Integers of any size, limited only by the available memory. See Section 6.2, “Type long: Extended-precision integers” (p. 8).</td>
<td>42L, -3L, 100000000000000L</td>
</tr>
<tr>
<td>bool</td>
<td>The two Boolean values True and False. See Section 6.3, “Type bool: Boolean truth values” (p. 8).</td>
<td>True, False</td>
</tr>
<tr>
<td>float</td>
<td>Floating-point numbers; see Section 6.4, “Type float: Floating-point numbers” (p. 9).</td>
<td>3.14159, -1.0, 6.0235e23</td>
</tr>
<tr>
<td>complex</td>
<td>Complex numbers. If the idea of computing with the square root of -1 bothers you, just ignore this type, otherwise see Section 6.5, “Type complex: Imaginary numbers” (p. 9).</td>
<td>(3.2+4.9j), (0+3.42e-3j)</td>
</tr>
<tr>
<td>str</td>
<td>Strings of 8-bit characters; see Section 7.2, “Type str: Strings of 8-bit characters” (p. 13). Strings can be empty: write such a string as “” or ‘’.</td>
<td>'Sir Robin', &quot;xyz&quot;, &quot;I'd've&quot;</td>
</tr>
<tr>
<td>unicode</td>
<td>Strings of 32-bit Unicode characters; see Section 7.3, “Type unicode: Strings of 32-bit characters” (p. 24).</td>
<td>u'Fred', u'\u03fa'</td>
</tr>
<tr>
<td>list</td>
<td>A mutable sequence of values; see Section 7.4, “Type list: Mutable sequences” (p. 25).</td>
<td>['dot', 'dash'];[]</td>
</tr>
<tr>
<td>tuple</td>
<td>An immutable sequence of values; see Section 7.5, “Type tuple: Immutable sequences” (p. 30).</td>
<td>('dot', 'dash');(); ('singleton',)</td>
</tr>
<tr>
<td>dict</td>
<td>Use dict values (dictionaries) to structure data as look-up tables; see Section 9, “Type dict: Dictionaries” (p. 36).</td>
<td>{'go':1, 'stop':2};{}</td>
</tr>
<tr>
<td>file</td>
<td>A file being read or written; see Section 10, “Type file: Input and output files” (p. 40).</td>
<td>open('/etc/motd')</td>
</tr>
<tr>
<td>None</td>
<td>A special, unique value that may be used where a value is required but there is no obvious value. See Section 11, “None: The special placeholder value” (p. 43).</td>
<td>None</td>
</tr>
</tbody>
</table>

6. Numeric types

Python has a number of different types used for representing numbers.

6.1. Type int: Integers

Python values of type int represent integers, that is, whole numbers in the range \([-2^{31}, 2^{31}-1]\), roughly plus or minus two billion.

You can represent a value in octal (base 8) by preceding it with a zero. Similarly, use a leading “0x” to represent a value in hexadecimal (base 16). Examples in conversational mode:
To convert other numbers or character strings to type `int`, see Section 13.16, “`int()`: Convert to `int` type” (p. 50).

If you perform operations on `int` values that result in numbers that are too large, Python automatically converts them to `long` type; see Section 6.2, “Type `long`: Extended-precision integers” (p. 8).

### 6.2. Type `long`: Extended-precision integers

Values of `long` type represent whole numbers, but they may have many more than the nine or ten digits allowed by `int` type. In practice, the number of digits in a `long` value is limited only by processor memory size.

To write a `long`-type constant, use the same syntax as for `int`-type constants, but place a letter `L` immediately after the last digit. Also, if an operation on `int` values results in a number too large to represent as an `int`, Python will automatically converted it to type `long`.

```python
>>> 100 * 100
10000
>>> 100L * 100L
100000L
>>> 1000000000*1000000000L
1000000000000000000L
>>> 0xffffL
65535L
```

To convert a value of a different numeric type or a string of characters to a `long` value, see Section 13.20, “`long()`: Convert to `long` type” (p. 51).

### 6.3. Type `bool`: Boolean truth values

A value of `bool` type represents a Boolean (true or false) value. There are only two values, written in Python as “`True`” and “`False`”.

Internally, `True` is represented as 1 and `False` as 0, and they can be used in numeric expressions as those values.

Here’s an example. In Python, the expression “`a < b`” compares two values `a` and `b`, and returns `True` if `a` is less than `b`, `False` is `a` is greater than or equal to `b`.

```python
>>> 2 < 3
True
>>> 3 < 2
False
>>> True+4
5
>>> False * False
0
```
These values are considered \texttt{False} wherever true/false values are expected, such as in an \texttt{if} statement:

- The \texttt{bool} value \texttt{False}.
- Any numeric zero: the \texttt{int} value \texttt{0}, the \texttt{float} value \texttt{0.0}, the \texttt{long} value \texttt{0L}, or the \texttt{complex} value \texttt{0j}.
- Any empty sequence: the \texttt{str} value \texttt{''}, the \texttt{unicode} value \texttt{u''}, the empty \texttt{list} value [], or the empty \texttt{tuple} value ()
- Any empty mapping, such as the empty \texttt{dict} (dictionary) value {}.
- The special value \texttt{None}.

All other values are considered \texttt{True}. To convert any value to a Boolean, see Section 13.4, “\texttt{bool()}: Convert to Boolean” (p. 46).

### 6.4. Type float: Floating-point numbers

Values of this type represent real numbers, with the usual limitations of IEEE-754 floating point type: it cannot represent very large or very small numbers, and the precision is limited to only about 15 digits. For complete details on the IEEE-754 standard and its limitations, see the Wikipedia article\footnote{http://en.wikipedia.org/wiki/IEEE_754-1985}.

A floating-point constant may be preceded by a “+” or “-” sign, followed by a string of one or more digits containing a decimal point (\texttt{.}).

For very large or small numbers, you may express the number in exponential notation by appending a letter “e” followed by a power of ten (which may be preceded by a sign).

For example, Avogadro’s Number gives the number of atoms of carbon in 12 grams of carbon\footnote{http://www.nist.gov/pml/data/avogadro.cdf}^12, and is written as $6.022148\times 10^{23}$. In Python that would be \texttt{6.022148e23}.

Please note that calculations involving \texttt{float} type are approximations. In calculator mode, Python will display the numbers to their full precision, so you may see a number that is very close to what you expect, but not exact.

```python
>>> 1.0/7.0
0.14285714285714285
>>> -2*-4.2e37
8.4000000000000004e+37
```

### 6.5. Type complex: Imaginary numbers

Mathematically, a complex number is a number of the form $A+Bi$ where $i$ is the imaginary number, equal to the square root of $-1$.

Complex numbers are quite commonly used in electrical engineering. In that field, however, because the symbol $i$ is used to represent current, they use the symbol $j$ for the square root of $-1$. Python adheres to this convention: a number followed by “j” is treated as an imaginary number. Python displays complex numbers in parentheses when they have a nonzero real part.

```python
>>> 5j
5j
>>> 1+2.56j
(1+2.5600000000000001j)
>>> (1+2.56j)*(-1-3.44j)
(7.8064-6j)
```

\footnote{http://en.wikipedia.org/wiki/IEEE_754-1985}
Unlike Python’s other numeric types, complex numbers are a composite quantity made of two parts: the real part and the imaginary part, both of which are represented internally as float values. You can retrieve the two components using attribute references. For a complex number \( C \):

- \( C\text{.real} \) is the real part.
- \( C\text{.imag} \) is the imaginary part as a float, not as a complex value.

```python
def main():
    a = (1+2.56j)*(-1-3.44j)
    print(a)
    print(a.real)
    print(a.imag)
```

To construct a complex value from two float values, see Section 13.7, “complex(): Convert to complex type” (p. 47).

## 7. Sequence types

The next four types described (str, unicode, list and tuple) are collectively referred to as sequence types.

Each sequence value represents an ordered set in the mathematical sense, that is, a collection of things in a specific order.

Python distinguishes between mutable and immutable sequences:

- An immutable sequence can be created or destroyed, but the number, sequence, and values of its elements cannot change.
- The values of a mutable sequence can be changed. Any element can be replaced or deleted, and new elements can be added at the beginning, the end, or in the middle.

There are four sequence types, but they share most of the same operations.

- Section 7.1, “Operations common to all the sequence types” (p. 10).
- Section 7.2, “Type str: Strings of 8-bit characters” (p. 13) (immutable).
- Section 7.3, “Type unicode: Strings of 32-bit characters” (p. 24) (immutable).
- Section 7.4, “Type list: Mutable sequences” (p. 25) (mutable).
- Section 7.5, “Type tuple: Immutable sequences” (p. 30) (immutable).

### 7.1. Operations common to all the sequence types

These functions work on values of the four sequence types: int, unicode, tuple, and list.

- Section 13.18, “\texttt{len()}: Number of elements” (p. 50).
- Section 13.22, “\texttt{max()}: Largest element of a sequence” (p. 51).
- Section 13.23, “\texttt{min()}: Smallest element of a sequence” (p. 52).

These operators apply to sequences.
**$S_1+S_2$**

Concatenation—for two sequences $S_1$ and $S_2$ of the same type, a new sequence containing all the elements from $S_1$ followed by all the elements of $S_2$.

```python
>>> "rat" + "bag"
'ratbag'
>>> [1, 2, 3] + [5, 7, 11, 13] + [15]
[1, 2, 3, 5, 7, 11, 13, 15]
>>> ('roy', 'g') + ('biv',)
('roy', 'g', 'biv')
```

**$S^n$**

For a sequence $S$ and a positive integer $n$, the result is a new sequence containing all the elements of $S$ repeated $n$ times.

```python
>>> 'worra'*8
'worraworraworraworraworraworraworrora'
>>> [0]*4
[0, 0, 0, 0]
>>> (True, False)*5
(True, False, True, False, True, False, True, False, True, False)
```

**$x$ in $S$**

Is any element of a sequence $S$ equal to $x$?

For convenience in searching for substrings, if the sequence to be searched is a string, the $x$ operand can be a multi-character string. In that case, the operation returns True if $x$ is found anywhere in $S$.

```python
>>> 1 in [2, 4, 6, 0, 8, 0]
False
>>> 0 in [2, 4, 6, 0, 8, 0]
True
>>> 'a' in 'banana'
True
>>> 3.0 in (2.5, 3.0, 3.5)
True
>>> "baz" in "rowrbazzle"
True
```

**$x$ not in $S$**

Are all the elements of a sequence $S$ not equal to $x$?

```python
>>> 'a' not in 'banana'
False
>>> 'x' not in 'banana'
True
```

**$S[i]$**

Subscripting: retrieve the $i$th element of $s$, counting from zero. If $i$ is greater than or equal to the number of elements of $S$, an IndexError exception is raised.

```python
>>> 'Perth'[0]
'P'
>>> 'Perth'[1]
```

New Mexico Tech Computer Center

Python 2.5 quick reference
S[i:j]
Slicing: For a sequence S and two integers i and j, return a new sequence with copies of the elements of S between positions i and j.

The values used in slicing refer to the positions between elements, where position zero is the position before the first element; position 1 is between the first and second element; and so on.

You can also specify positions relative to the end of a sequence. Position -1 is the position before the last element; -2 is the position before the second-to-last element; and so on.

You can omit the starting position to obtain a slice starting at the beginning. You can omit the ending position to get all the elements through the last.

For example, here is a diagram showing three slices of the string 'abcdef'.

>>> 'abcdef'[2:5]
'cde'
>>> 'abcdef'[:3]
'abc'
>>> 'abcdef'[3:]
'def'
>>> (90, 91, 92, 93, 94, 95)[2:5]
(92, 93, 94)

S[i:j:k]
You can use a slice expression like this to select every kth element. Examples:

>>> teens = range(13, 20)
>>> teens
[13, 14, 15, 16, 17, 18, 19]
>>> teens[::2]
[13, 15, 17, 19]
>>> teens[1::2]
[14, 16, 18]
>>> teens[1:5]
7.2. Type str: Strings of 8-bit characters

Python has two string types. Type `str` holds strings of zero or more 8-bit characters, while `unicode` strings provide full support of the expanded Unicode character set; see Section 7.3, “Type unicode: Strings of 32-bit characters” (p. 24).

In addition to the functions described in Section 7.1, “Operations common to all the sequence types” (p. 10), these functions apply to strings:

- Section 13.29, “raw_input(): Prompt and read a string from the user” (p. 53).
- Section 13.35, “str(): Convert to str type” (p. 56).
- Section 13.40, “unicode(): Convert to a Unicode string” (p. 58).

7.2.1. String constants

There are many forms for string constants:

- `'...'`: You may enclose the string in single-quotes.
- `"..."`: You may instead enclose it in double-quotes. Internally, there is absolutely no difference. To include a double-quote character inside the string, use the escape sequence `"\"`.

In conversational mode, Python will generally display values using single-quotes. If the string contains single-quotes but no double-quotes, Python will display the value using double-quotes. If the string contains both, the value will be displayed in single-quotes, with single-quote characters inside the value displayed as the escape sequence `\'`.

- `'...'...'''`: You may enclose your string between three single quotes in a row. The difference is that you can continue such a string over multiple lines, and the line breaks will be included in the string as newline characters.
- `"""..."""`: You can use three sets of double quotes. As with three sets of single quotes, line breaks are allowed and preserved as `\n` characters. If you use these triply-quoted strings in conversational mode, continuation lines will prompt you with `...`.

```python
>>> 'Penguin'
'Penguin'
>>> "ha\'penny"
"ha\'penny"
>>> "Single \' and double\" quotes"
'Single \' and double" quotes'
>>> \
...''
```

```python
>>> "Single \' and double\" quotes"
'Single \' and double" quotes'
```
In addition, you can use any of these escape sequences inside a string constant:

### Table 2. String escape sequences

<table>
<thead>
<tr>
<th>Escape Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>\n</code></td>
<td>A backslash at the end of a line is ignored.</td>
</tr>
<tr>
<td><code>\&quot;</code></td>
<td>Backslash ()</td>
</tr>
<tr>
<td><code>\'</code></td>
<td>Closing single quote (')</td>
</tr>
<tr>
<td><code>[tab]</code></td>
<td>Double-quote character (&quot;)</td>
</tr>
<tr>
<td><code>\r</code></td>
<td>Newline (ASCII LF or linefeed)</td>
</tr>
<tr>
<td><code>\b</code></td>
<td>Backspace (in ASCII, the BS character)</td>
</tr>
<tr>
<td><code>\f</code></td>
<td>Formfeed (ASCII FF)</td>
</tr>
<tr>
<td><code>\r</code></td>
<td>Carriage return (ASCII CR)</td>
</tr>
<tr>
<td><code>\t</code></td>
<td>Horizontal tab (ASCII HT)</td>
</tr>
<tr>
<td><code>\v</code></td>
<td>Vertical tab (ASCII VT)</td>
</tr>
<tr>
<td><code>\000</code></td>
<td>The character with octal code <code>000</code>, e.g., <code>'\177'</code>.</td>
</tr>
<tr>
<td><code>\xhh</code></td>
<td>The character with hexadecimal value <code>hh</code>, e.g., <code>'\xFF'</code>.</td>
</tr>
</tbody>
</table>

**Raw strings:** If you need to use a lot of backslashes inside a string constant, and doubling them is too confusing, you can prefix any string with the letter r to suppress the interpretation of escape sequences. For example, `'\\' contains two backslashes, but `r'\\' contains four. Raw strings are particularly useful with Section 21.5, “re: Regular expression pattern-matching” (p. 117).

### 7.2.2. The string format operator

In addition to the common sequence operations described in Section 7.1, “Operations common to all the sequence types” (p. 10), strings support the format operator “%”. This extremely useful operator uses a template string to specify how to format one or more values. Here is the general form:

```
f %v
```

where f is the template string and v specifies the value or values to be formatted using that template. If multiple values are to be formatted, v must be a tuple.

The template string may contain any mixture of ordinary text and format codes. A format code always starts with a percent (%) symbol. See Table 4, “Format codes” (p. 15).

The result of a format operation consists of the ordinary characters from the template with values substituted within them wherever a format code occurs. A conversational example:

```
>>> print "We have %d pallets of %s today." % (49, "kiwis")
We have 49 pallets of kiwis today.
```

In the above example, there are two format codes. Code “%d” means “substitute a decimal number here,” and code “%s” means “substitute a string value here”. The number 49 is substituted for the first format code, and the string "kiwis" replaces the second format code.

In general, format codes have this form:

```
%[p][m[n]c
```
Table 3. Parts of the format operator

<table>
<thead>
<tr>
<th>p</th>
<th>An optional prefix; see Table 5, “Format code prefixes” (p. 16).</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>Specifies the total desired field width. The result will never be shorter than this value, but may be longer if the value doesn't fit; so, &quot;%5d&quot; % 1234 yields &quot; 1234&quot;, but &quot;%2d&quot; % 1234 yields &quot;1234&quot;. If the value is negative, values are left-aligned in the field whenever they don't fill the entire width.</td>
</tr>
<tr>
<td>n</td>
<td>For float values, this specifies the number of digits after the decimal point.</td>
</tr>
<tr>
<td>c</td>
<td>Indicates the type of formatting.</td>
</tr>
</tbody>
</table>

Here are the format type codes, c in the general expression above:

Table 4. Format codes

<table>
<thead>
<tr>
<th>Format code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%s</td>
<td>Format a string. For example, '%-3s' % 'xy' yields 'xy '; the width (-3) forces left alignment.</td>
</tr>
<tr>
<td>%d</td>
<td>Decimal conversion. For example, '%3d' % -4 yields the string ' -4'.</td>
</tr>
<tr>
<td>%e</td>
<td>Exponential format; allow four characters for the exponent. Examples: '%08.1e' % 1.9783 yields '0002.0e+00'.</td>
</tr>
<tr>
<td>%E</td>
<td>Same as %e, but the exponent is shown as an uppercase E.</td>
</tr>
<tr>
<td>%f</td>
<td>For float type. E.g., '%4.1f' % 1.9783 yields ' 2.0'.</td>
</tr>
<tr>
<td>%g</td>
<td>General numeric format. Use %f if it fits, otherwise use %e.</td>
</tr>
<tr>
<td>%o</td>
<td>Same as %G, but an uppercase E is used for the exponent if there is one.</td>
</tr>
<tr>
<td>%o</td>
<td>Octal (base 8). For example, '%o' % 13 yields ' 15'.</td>
</tr>
<tr>
<td>%x</td>
<td>Hexadecimal (base 16). For example, '%x' % 247 yields ' f7'.</td>
</tr>
<tr>
<td>%X</td>
<td>Same as %x, but capital letters are used for the digits A-F. For example, '%04X' % 247 yields '00F7'; the leading zero in the length (04) requests that Python fill up any empty leading positions with zeroes.</td>
</tr>
<tr>
<td>%c</td>
<td>Convert an integer to the corresponding ASCII code. For example, '%c' % 0x61 yields the string 'a'.</td>
</tr>
<tr>
<td>%%</td>
<td>Places a percent sign (%) in the result. Does not require a corresponding value. Example: &quot;Energy at %d%%.&quot; % 88 yields the value 'Energy at 88%.'</td>
</tr>
</tbody>
</table>
Table 5. Format code prefixes

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>For numeric types, forces the sign to appear even for positive values.</td>
</tr>
<tr>
<td>-</td>
<td>Left-justifies the value in the field.</td>
</tr>
<tr>
<td>0</td>
<td>For numeric types, use zero fill. For example, ‘%04d’ % 2 produces the value ‘0002’.</td>
</tr>
<tr>
<td>#</td>
<td>With the %0 (octal) format, append a leading &quot;0&quot;; with the %x (hexadecimal) format, append a leading &quot;0x&quot;; with the %g (general numeric) format, append all trailing zeroes. Examples:</td>
</tr>
</tbody>
</table>

```python
>>> '%4o' % 127
' 177'
>>> '%#4o' % 127
'0177'
>>> '%x' % 127
'7f'
>>> '%#x' % 127
'0x7f'
>>> '%10.5g' % 0.5
' 0.5'
>>> '%#10.5g' % 0.5
' 0.50000'
```

7.2.3. String formatting from a dictionary

You can use the string format operator % to format a set of values from a dictionary \( D \) (see Section 9, “Type dict: Dictionaries” (p. 36)):

\[
f \% D
\]

In this form, the general form for a format code is:

\[
%(k)[p][m[.n]]c
\]

where \( k \) is a key in dictionary \( D \), and the rest of the format code is as in the usual string format operator. For each format code, the value of \( D[k] \) is used. Example:

```python
>>> named = {‘last’: ‘Poe’, ‘first’: ‘Aloysius’}
>>> ‘Dear %(first)s %(last)s:’ % named
‘Dear Aloysius Poe:
```

7.2.4. Definition of “whitespace”

In Python, these characters are considered whitespace:

<table>
<thead>
<tr>
<th>Escape sequence</th>
<th>ASCII(^6) name</th>
<th>English name</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘ ‘</td>
<td>SP</td>
<td>space</td>
</tr>
<tr>
<td>‘\n’</td>
<td>NL</td>
<td>newline</td>
</tr>
<tr>
<td>‘\r’</td>
<td>CR</td>
<td>carriage return</td>
</tr>
<tr>
<td>‘\t’</td>
<td>HT</td>
<td>horizontal tab</td>
</tr>
</tbody>
</table>

\(^6\) http://en.wikipedia.org/wiki/ASCII
<table>
<thead>
<tr>
<th>Escape sequence</th>
<th>ASCII name</th>
<th>English name</th>
</tr>
</thead>
<tbody>
<tr>
<td>'\f'</td>
<td>FF</td>
<td>form feed</td>
</tr>
<tr>
<td>'\v'</td>
<td>VT</td>
<td>vertical tab</td>
</tr>
</tbody>
</table>

### 7.2.5. Methods on str values

These methods are available on any string value `S`.

**`S.capitalize()`**
Return `S` with its first character capitalized (if a letter).

```python
g>>> 'e e cummings'.capitalize()
'E e cummings'
g>>> '---abc---'.capitalize()
'---abc---'
```

**`S.center(w)`**
Return `S` centered in a string of width `w`, padded with spaces. If `w<=len(S)`, the result is a copy of `S`. If the number of spaces of padding is odd, the extra space will placed after the centered value.
Example:

```python
g>>> 'x'.center(4)
'  x  '
```

**`S.count(t[, start[, end]])`**
Return the number of times string `t` occurs in `S`. To search only a slice `S[start:end]` of `S`, supply `start` and `end` arguments.

```python
g>>> 'banana'.count('a')
3
g>>> 'bananana'.count('na')
3
g>>> 'banana'.count('a', 3)
2
g>>> 'banana'.count('a', 3, 5)
1
```

**`S.decode( encoding )`**
If `S` contains an encoded Unicode string, this method will return the corresponding value as `unicode` type. The `encoding` argument specifies which decoder to use; typically this will be the string `'utf_8'` for the UTF-8 encoding. For discussion and examples, see Section 7.3.1, “The UTF-8 encoding” (p. 24).

**`S.endswith(t[, start[, end]])`**
Predicate to test whether `S` ends with string `t`. If you supply the optional `start` and `end` arguments, it tests whether the slice `S[start:end]` ends with `t`.

```python
g>>> 'bishop'.endswith('shop')
True
g>>> 'bishop'.endswith('bath and wells')
False
g>>> 'bishop'[3:5]
```

---

'ho'
>>> 'bishop'.endswith('o', 3, 5)
True

S.expandtabs([tabsize])
Returns a copy of S with all tabs replaced by one or more spaces. Each tab is interpreted as a request to move to the next “tab stop”. The optional tabsize argument specifies the number of spaces between tab stops; the default is 8.

Here is how the function actually works. The characters of S are copied to a new string T one at a time. If the character is a tab, it is replaced by enough tabs so the new length of T is a multiple of the tab size (but always at least one space).

>>> 'X\tY\tZ'.expandtabs()
'X Y Z'
>>> 'X\tY\tZ'.expandtabs(4)
'X Y Z'
>>> 'a\ttb\ttccc\ttddd\ttteeee\ttfffff'.expandtabs(4)
'a bb ccc dddd eeee fffff'

S.find(t[,start[,end]])
If string t is not found in S, return -1; otherwise return the index of the first position in S that matches t.

The optional start and end arguments restrict the search to slice S[start:end].

>>> 'banana'.find('an')
1
>>> 'banana'.find('ape')
-1
>>> 'banana'.find('n', 3)
4
>>> 'council'.find('c', 1, 4)
-1

S.index(t[,start[,end]])
Works like .find(), but if t is not found, it raises a ValueError exception.

>>> 'council'.index('co')
0
>>> 'council'.index('phd')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: substring not found

S.isalnum()
Predicate that tests whether S is nonempty and all its characters are alphanumeric.

>>> ''.isalnum()
False
>>> 'abc123'.isalnum()
True
>>> '&*$#&*()abc123'.isalnum()
False
**S.isalpha()**
Predicate that tests whether S is nonempty and all its characters are letters.

```python
>>> 'abc123'.isalpha()
False
>>> 'MaryRecruiting'.isalpha()
True
>>> ''.isalpha()
False
```

**S.isdigit()**
Predicate that tests whether S is nonempty and all its characters are digits.

```python
>>> 'abc123'.isdigit()
False
>>> ''.isdigit()
False
>>> '2415'.isdigit()
True
```

**S.islower()**
Predicate that tests whether S is nonempty and all its letters are lowercase (non-letter characters are ignored).

```python
>>> ''.islower()
False
>>> 'abc123'.islower()
True
>>> 'ABC123'.islower()
False
```

**S.isspace()**
Predicate that tests whether S is nonempty and all its characters are whitespace characters.

```python
>>> ''.isspace()
False
>>> '\t\n\r'.isspace()
True
>>> 'killer \t \n rabbit'.isspace()
False
```

**S.istitle()**
A predicate that tests whether S has “title case”. In a title-cased string, uppercase characters may appear only at the beginning of the string or after some character that is not a letter. Lowercase characters may appear only after an uppercase letter.

```python
>>> 'abc def GHI'.istitle()
False
>>> 'Abc Def Ghi'.istitle()
True
```

**S.isupper()**
Predicate that tests whether S is nonempty and all its letters are uppercase letters (non-letter characters are ignored).
S.join(L)
L must be a sequence of strings. The returned value is a string containing the members of the sequence with copies of the delimiter string S inserted between them.

One quite common operation is to use the empty string as the delimiter to concatenate the elements of a sequence.

Examples:

>>> '/'.join(['never', 'pay', 'plan'])
'never/pay/plan'

>>> '(***)'.join ( ('Property', 'of', 'the', 'zoo') )
'Property(***)of(***)the(***)zoo'

>>> '.join(['anti', 'dis', 'establish', 'ment', 'arian', 'ism'])
'antidisestablishmentarianism'

S.ljust(w)
Return a copy of S left-justified in a field of width w, padded with spaces. If w<=len(S), the result is a copy of S.

>>> 'Ni'.ljust(4)
'Ni '

S.lower()
Returns a copy of S with all uppercase letters replaced by their lowercase equivalent.

>>> "I like SHOUTING!".lower()
'i like shouting!

S.lstrip([c])
Return S with all leading characters from string c removed. The default value for c is a string containing all the whitespace characters.

>>> '    Run    

away!

'.lstrip()
'Run    

away!

'

>>> "***Done***".lstrip('*')
'Done***

>>> "(*)(*)(*Undone*)".lstrip ( "(*)" )
'Undone*)

S.partition(d)
Searches string S for the first occurrence of some delimiter string d. If S contains the delimiter, it returns a tuple (pre, d, post), where pre is the part of S before the delimiter, d is the delimiter itself, and post is the part of S after the delimiter.

If the delimiter is not found, this method returns a 3-tuple (S, '', '').

>>> "Daffy English kniggets!".partition(' ') ('Daffy', ' ', 'English kniggets!')
S.replace(old,new[,max])
Return a copy of S with all occurrences of string old replaced by string new. Normally, all occurrences are replaced; if you want to limit the number of replacements, pass that limit as the max argument.

>>> 'Frenetic'.replace('e', 'x')
'Frxnxtic'
>>> 'Frenetic'.replace('e', '###')
'Fr###n###tic'
>>> 'banana'.replace('an', 'erzerk')
'berzerkerzerka'
>>> 'banana'.replace('a', 'x', 2)
'bxnxna'

S.rfind(t[,start[,end]])
Like .find(), but if t occurs in S, this method returns the highest starting index.

>>> 'banana'.find('a')
1
>>> 'banana'.rfind('a')
5

S.rindex(t[,start[,end]])
Similar to S.index(), but it returns the last index in S where string t is found. It will raise a ValueError exception if the string is not found.

>>> "Just a flesh wound.".index('s')
2
>>> "Just a flesh wound.".rindex('s')
10
>>> "Just a flesh wound.".rindex('xx')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: substring not found

S.rjust(w[,fill])
Return a copy of S right-justified in a field of width w, padded with spaces. If w<=len(S), the result is a copy of S.

To pad values with some character other than a space, pass that character as the optional second argument.

>>> '123'.rjust(5)
' 123'
>>> '123'.rjust(5, '*')
'***123'

S.rpartition(d)
Similar to S.partition(), except that it finds the last occurrence of the delimiter.
>>> "Daffy English kniggets!".rpartition( ' ')
('Daffy English', ' ', 'kniggets!')

>>> "a*b***c*d".rpartition("***")
('a*b*', '***', 'c*d')

\[ S.rsplit(d[,max]) \]
Similar to \[ S.split(d[,max]) \], except that if there are more fields than \[ max \], the split fields are taken from the end of the string instead of from the beginning.

```python
>>> "I am Zoot's identical twin sister, Dingo.".rsplit(None, 2)
['"I am Zoot's identical twin", 'sister,', 'Dingo.']
```

\[ S.rstrip([c]) \]
Return \[ S \] with all trailing characters from string \[ c \] removed. The default value for \[ c \] is a string containing all the whitespace characters.

```python
>>> ' \t \n Run \t \n away ! \n \t '.rstrip()
' \t \n Run \t \n away !'
```

\[ S.split([d[,max]]) \]
Returns a list of strings \[ [s_0, s_1, ...] \] made by splitting \[ S \] into pieces wherever the delimiter string \[ d \] is found. The default is to split up \[ S \] into pieces wherever clumps of one or more whitespace characters are found.

```python
>>> "I'd annex \t \r the Sudetenland".split()
["I'd", 'annex', 'the', 'Sudetenland']
>>> '3/crunchy frog/ Bath & Wells'.split('/')
['3', 'crunchy frog', ' Bath & Wells']
>>> '//Norwegian Blue/'.split('/')
['', '', 'Norwegian Blue', '']
>>> 'never<*>pay<*>plan<*>'.split('<*>')
['never', 'pay', 'plan', '']
```

The optional \[ max \] argument limits the number of pieces removed from the front of \[ S \]. The resulting list will have no more than \[ max+1 \] elements.

To use the \[ max \] argument while splitting the string on clumps of whitespace, pass \[ None \] as the first argument.

```python
>>> 'a/b/c/d/e'.split('/', 2)
['a', 'b', 'c/d/e']
>>> 'a/b'.split('/', 2)
['a', 'b']
```

\[ S.splitlines([keepends]) \]
Splits \[ S \] into lines and returns a list of the lines as strings. Discards the line separators unless the optional \[ keepends \] arguments is true.

```python
>>> """Is that
... an ocarina?""".splitlines()
['"Is that', ' an ocarina?']
```
... Sir Robin of Camelot.""".splitlines(True)
['What is your name?
', 'Sir Robin of Camelot.']

S.startswith(t[,start[,end]])
Predicate to test whether S starts with string t. Otherwise similar to .endswith().

>>> "bishop".startswith('bish')
True
>>> "bishop".startswith('The')
False

S.strip([c])
Return S with all leading and trailing characters from string c removed. The default value for c is a string containing all the whitespace characters.

>>> ' 	 \n Run \t \n away ! \n 	 '.strip()
'Run \t \n away !'

S.swapcase()
Return a copy of S with each lowercase character replaced by its uppercase equivalent, and vice versa.

>>> "abcDEF".swapcase()
'ABCdef'

S.title()
Returns the characters of S, except that the first letter of each word is uppercased, and other letters are lowercased.

>>> "huge...tracts of land".title()
'Huge...Tracts Of Land'

S.translate(new[,drop])
This function is used to translate or remove each character of S. The new argument is a string of exactly 256 characters, and each character x of the result is replaced by new[ord(x)].

If you would like certain characters removed from S before the translation, provide a string of those characters as the drop argument.

For your convenience in building the special 256-character strings used here, see the definition of the maketrans() function of Section 21.2, “string: Utility functions for strings” (p. 113), where you will find examples.

S.upper()
Return a copy of S with all lowercase characters replaced by their uppercase equivalents.

>>> 'I like shouting'.upper()
'I LIKE SHOUTING'

S.zfill(w)
Return a copy of S left-filled with '0' characters to width w.

>>> '12'.zfill(9)
'000000012'
7.3. Type unicode: Strings of 32-bit characters

With the advent of the Web as medium for worldwide information interchange, the Unicode character set has become vital. For general background on this character set, see the Unicode homepage[^1].

To get a Unicode string, prefix the string with u. For example:

```python
u'klarn'
```

is a five-character Unicode string.

To include one of the special Unicode characters in a string constant, use these escape sequences:

<table>
<thead>
<tr>
<th>Escape Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\xHH</td>
<td>For a code with the 8-bit hexadecimal value HH.</td>
</tr>
<tr>
<td>\uHHHH</td>
<td>For a code with the 16-bit hexadecimal value HHHH.</td>
</tr>
<tr>
<td>\UHHHHHHHHH</td>
<td>For a code with the 32-bit hexadecimal value HHHHHHHH.</td>
</tr>
</tbody>
</table>

Examples:

```python
>>> u'Klarn.'
u'Klarn.'

>>> u'Non-breaking-\xa0-space.'
u'Non-breaking-\xa0-space.'

>>> u'Less-than-or-equal symbol: \u2264'
u'Less-than-or-equal symbol: \u2264'

>>> u'Phoenician letter 'wau': \U00010905'
u'Phoenician letter 'wau': \U00010905'

>>> len(u'\U00010905')
1
```

All the operators and methods of `str` type are available with `unicode` values.

Additionally, for a Unicode value `U`, use this method to encode its value as a string of type `str`:

```python
U.encode( encoding )
```

Return the value of `U` as type `str`. The `encoding` argument is a string that specifies the encoding method. In most cases, this will be `\'utf_8\'`. For discussion and examples, see Section 7.3.1, “The UTF-8 encoding” (p. 24).

7.3.1. The UTF-8 encoding

How, you might ask, do we pack 32-bit Unicode characters into 8-bit bytes? Quite prevalent on the Web and the Internet generally is the UTF-8 encoding, which allows any of the Unicode characters to be represented as a string of one or more 8-bit bytes.

First, some definitions:

- A **code point** is a number representing a unique member of the Unicode character set.
- The Unicode code points are visualized as a three-dimensional structure made of **planes**, each of which has a range of 65536 code points organized as 256 **rows** of 256 **columns** each.

The low-order eight bits of the code point select the column; the next eight more significant bits select the row; and the remaining most significant bits select the plane.

[^1]: http://www.unicode.org/
This diagram shows how UTF-8 encoding works. The first 128 code points (hexadecimal 00 through 7F) are encoded as in normal 7-bit ASCII, with the high-order bit always 0. For code points above hex 7F, all bytes have the high-order (0x80) bit set, and the bits of the code point are distributed through two, three, or four bytes, depending on the number of bits needed to represent the code point value.

<table>
<thead>
<tr>
<th>Code point</th>
<th>UTF-8 encoded</th>
</tr>
</thead>
<tbody>
<tr>
<td>7x</td>
<td>0 x x x x x x</td>
</tr>
<tr>
<td>5y 6z</td>
<td>1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>4x 6y 6z</td>
<td>1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>3w 6x 6y 6z</td>
<td>1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

To encode a Unicode string \( U \), use this method:

\[ U . e n c o d e ( ' u t f _ 8' ) \]

To decode a regular \( s t r \) value \( S \) that contains a UTF-8 encoded value, use this method:

\[ S . d e c o d e ( ' u t f _ 8' ) \]

Examples:

```python
>>> tilde='~
>>> tilde.encode('utf_8')
'~'
>>> u16 = u'\u0456'
>>> s = u16.encode('utf_8')
>>> s
'\xd1\x96'
>>> s.decode('utf_8')
u'\u0456'
>>> u32 = u'\U000E1234'
>>> s = u32.encode('utf_8')
>>> s
'\xf3\x8a\x88\xb4'
>>> s.decode('utf_8')
u'\U000e1234'
```

UTF-8 is not the only encoding method. For more details, consult the documentation for the Python module `codecs`.8

### 7.4. Type list: Mutable sequences

To form values into a sequence, use Python's `list` type if you are going to change, delete, or add values to the sequence. For a discussion of when to use `list` and when to use `tuple`, see Section 7.5, “Type tuple: Immutable sequences” (p. 30).

To create a list, enclose a list of zero or more comma-separated values inside square brackets, “[...].”

Examples:

8 [http://docs.python.org/lib/module-codecs.html](http://docs.python.org/lib/module-codecs.html)
You can also create a list by performing specific operations on each element of some sequence; see Section 7.4.2, “List comprehensions” (p. 29).

Lists support all the operations described under Section 7.1, “Operations common to all the sequence types” (p. 10). Methods available on lists are discussed in Section 7.4.1, “Methods on lists” (p. 26).

There are a number of functions that can be used with lists as well:

- Section 13.2, “all()”: Are all the elements of a sequence true? (p. 45).
- Section 13.3, “any()”: Are any of the members of a sequence true? (p. 45).
- Section 13.6, “cmp()”: Compare two values (p. 46).
- Section 13.10, “enumerate()”: Step through indices and values of a sequence (p. 48).
- Section 13.12, “filter()”: Extract qualifying elements from a sequence (p. 48).
- Section 13.17, “iter()”: Produce an iterator over a sequence (p. 50).
- Section 13.18, “len()”: Number of elements (p. 50).
- Section 13.19, “list()”: Convert to a list (p. 51).
- Section 13.21, “map()”: Apply a function to each element of a sequence (p. 51).
- Section 13.22, “max()”: Largest element of a sequence (p. 51).
- Section 13.23, “min()”: Smallest element of a sequence (p. 52).
- Section 13.28, “range()”: Generate an arithmetic progression as a list (p. 53).
- Section 13.30, “reduce()”: Sequence reduction (p. 54).
- Section 13.31, “reversed()”: Produce a reverse iterator (p. 54).
- Section 13.34, “xrange()”: Arithmetic progression generator (p. 58).
- Section 13.36, “sum()”: Total the elements of a sequence (p. 57).
- Section 13.41, “sorted()”: Sort a sequence (p. 56).
- Section 13.42, “zip()”: Combine multiple sequences (p. 59).

### 7.4.1. Methods on lists

For any list value $L$, these methods are available.

$L$.append($x$)

Append a new element $x$ to the end of list $L$. Does not return a value.

```python
>>> colors = ['red', 'green', 'blue']
>>> colors.append('indigo')
>>> colors
['red', 'green', 'blue', 'indigo']
```

$L$.count($x$)

Return the number of elements of $L$ that compare equal to $x$.

```python
>>> [59, 0, 0, 63, 0, 0].count(0)
5
>>> ['x', 'y'].count('Fomalhaut')
0
```

$L$.extend($S$)

Append another sequence $S$ to $L$.
>>> colors
['red', 'green', 'blue', 'indigo']
>>> colors.extend(['violet', 'pale puce'])
>>> colors
['red', 'green', 'blue', 'indigo', 'violet', 'pale puce']

```
L.index(x[, start[, end]])
```
If \(L\) contains any elements that equal \(x\), return the index of the first such element, otherwise raise a `ValueError` exception.

The optional `start` and `end` arguments can be used to search only positions within the slice \(L[start:end]\).

```
>>> colors
['red', 'green', 'blue', 'indigo', 'violet', 'pale puce']
>>> colors.index('blue')
2
>>> colors.index('taupe')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: list.index(x): x not in list
```

```
L.insert(i, x)
```
Insert a new element \(x\) into list \(L\) just before the \(i\)th element, shifting all higher-number elements to the right. No value is returned.

```
>>> colors
['red', 'green', 'blue', 'indigo', 'violet', 'pale puce']
>>> colors[1]
'green'
>>> colors.insert(1, "yellow")
>>> colors
['red', 'yellow', 'green', 'blue', 'indigo', 'violet', 'pale puce']
```

```
L.pop([i])
```
Remove and return the element with index \(i\) from \(L\). The default value for \(i\) is -1, so if you pass no argument, the last element is removed.

```
>>> colors
['red', 'yellow', 'green', 'blue', 'indigo', 'violet', 'pale puce']
>>> tos = colors.pop()
>>> tos
'pale puce'
>>> colors
['red', 'yellow', 'green', 'blue', 'indigo', 'violet']
>>> colors[4]
```
>>> dye = colors.pop(4)
>>> dye
'indigo'
>>> colors
['red', 'yellow', 'green', 'blue', 'violet']

$L$.remove($x$)
Remove the first element of $L$ that is equal to $x$. If there aren't any such elements, raises ValueError.

>>> colors
['red', 'yellow', 'green', 'blue', 'violet']
>>> colors.remove('yellow')
>>> colors
['red', 'green', 'blue', 'violet']
>>> colors.remove('cornflower')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: list.remove(x): x not in list

>>> notMuch = [0, 0, 3, 0]
>>> notMuch.remove(0)
>>> notMuch
[0, 3, 0]
>>> notMuch.remove(0)
>>> notMuch
[3, 0]
>>> notMuch.remove(0)
>>> notMuch
[3]
>>> notMuch.remove(0)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: list.remove(x): x not in list

$L$.reverse()
Reverse the elements of $L$ in place. Does not return a result. Compare Section 13.31, “reversed(): Produce a reverse iterator” (p. 54).

>>> colors
['red', 'green', 'blue', 'violet']
>>> colors.reverse()
>>> colors
['violet', 'blue', 'green', 'red']

$L$.sort(cmp[, key[, reverse]])
Sort list $L$ in place. Does not return a result. Compare Section 13.34, “sorted(): Sort a sequence” (p. 56).

The reordering is guaranteed to be stable—that is, if two elements are considered equal, their order after sorting will not change.

While sorting, Python will use the built-in cmp() function to compare elements; see Section 13.6, “cmp(): Compare two values” (p. 46). You may provide, as the first argument to the .sort() method, your own comparator function to compare elements. This function must have the same calling sequence and return value convention as the built-in cmp() function: it must take two argu-
ments, and return a negative number of the first argument precedes the second, a positive number if the second argument precedes the first, or zero if they are considered equal.

You may also provide a “key extractor function” that is applied to each element to determine its key. This function must take one argument and return the value to be used as the sort key. If you want to provide a key extractor function but not a comparator function, pass None as the first argument to the method.

Additionally, you may provide a third argument of True to sort the sequence in descending order; the default behavior is to sort into ascending order.

```python
>>> temps=[67, 73, 85, 93, 92, 78, 95, 100, 104]
>>> temps.sort()
>>> temps
[67, 73, 78, 85, 92, 93, 95, 100, 104]
>>> def reverser(n1, n2):
...     '''Comparison function to use reverse order.'''
...     ...     return cmp(n2, n1)
... >>> temps.sort(reverser)
>>> temps
[104, 100, 95, 93, 92, 85, 78, 73, 67]
>>> def unitsDigit(n):
...     '''Returns only the units digit of n.'''
...     ...     return n % 10
... >>> temps.sort(None, unitsDigit)
>>> temps
[100, 92, 93, 73, 104, 95, 85, 67, 78]
>>> temps.sort(None, None, True)
>>> temps
[104, 100, 95, 93, 92, 85, 78, 73, 67]
```

### 7.4.2. List comprehensions

You can use a form called a list comprehension to create a list. The general form is:

```
[e
  for v_1 in s_1
  for v_2 in s_2
  ... if c]
```

where e is some expression, followed by one or more for clauses, optionally followed by an if clause.

The result is a list containing all the values of expression e after all the nested for loops have been run; the for loops have the same structure as in Section 16.4, “The for statement: Iteration over a sequence” (p. 79). If there is an “if” clause, it determines which values of e are added to the list: if the if condition is true, the value is added, otherwise it is not added.

This is perhaps easiest to explain with a few examples. In the first example, we construct a list containing the cubes of the numbers from 1 to 10, inclusive. The for loop generates the numbers 1, 2, ..., 10, and then the expression “x**3” cubes each one and appends it to the resulting list.
In the next example, we use two for loops. The outer loop generates the sequence [1, 2, 3], and the inner loop generates the sequence [50, 51]. The expression “x*1000 + y” is computed for each of the resulting six value sets for x and y, and the result is appended to the list being built.

```python
>>> [ x*1000 + y
...   for x in range(1,4)
...   for y in range(50, 52) ]
[1050, 1051, 2050, 2051, 3050, 3051]
```

In the next example, there are two nested loops, each generating the sequence [0, 1, 2]. For each of the nine trips through the inner loop, we test the values of x and y and discard the cases where they are equal. The expression “(y, x)” combines the two values into a 2-tuple.

```python
>>> [ (y, x)
...   for y in range(3)
...   for x in range(3)
...   if x != y ]
[(0, 1), (0, 2), (1, 0), (1, 2), (2, 0), (2, 1)]
```

## 7.5. Type tuple: Immutable sequences

For representing a sequence of values, Python has two similar container types: list and tuple.

An instance of a container type is basically a value that has other values inside it; we call the contained values elements.

So, when should you use a list, and when a tuple? In many contexts, either will work. However, there are important differences.

- Values of type list are mutable; that is, you can delete or add elements, or change the value of any of the elements inside the list.

Lists cannot be used in certain contexts. For example, you can’t use a list as the right-hand argument of the string format operator, or as the key in a dictionary.

```python
>>> "We have %d braces of %s." % (14, 'velociraptors')
'We have 14 braces of velociraptors.'
>>> "We have %d braces of %s." % [14, 'velociraptors']
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
TypeError: int argument required
>>> d={}
>>> d[(23,59)] = 'hike'
>>> d[(46,19)] = 'hut'
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
TypeError: list objects are unhashable
```

- Values of type tuple are immutable. Once you have assembled a tuple, you cannot add or delete elements or change the value of an element inside the tuple.
Among the reasons to use a tuple instead of a list:

- Tuples are allowed in certain contexts where lists are not, such as the right-hand argument of the string format operator, or as a key in a dictionary.
- If your program is in danger of running out of processor memory, tuples are slightly more efficient in their memory usage.

Write a literal tuple as a sequence of values in parentheses, separated by commas. There may be any number of values of any type or any mixture of types. There may be zero values.

To write a tuple with exactly one value, you must use this special syntax:

\[
(value,)
\]

That is, you must provide a comma before the closing “)”, in order to show that it is a tuple, and not just a parenthesized expression. Note especially the last two examples below:

```python
>>> ()
()
>>> ('farcical', 'aquatic', 'ceremony')
('farcical', 'aquatic', 'ceremony')
>>> ('Ni',)
('Ni',)
>>> ('Ni')
'Ni'
```

You may also convert an existing sequence of another type using Section 13.37, “tuple(): Convert to a tuple” (p. 57).

The tuple type does not have comprehensions (see Section 7.4.2, “List comprehensions” (p. 29)), but you can get the equivalent by applying the `tuple()` function to a list comprehension. Here is an example:

```python
>>> tuple([x**2 for x in (2,3,8)])
(4, 9, 64)
```

---

### 8. Types set and frozenset: Set types

Mathematically speaking, a set is an unordered collection of zero or more distinct elements. Python has two set types that represent this mathematical abstraction. Use these types when you care only about whether something is a member of the set or not, and you don't need them to be in any specific order.

The elements of a Python set must be immutable. In particular, you can't have list or dictionary elements in a set.

Most operations on sets work with both `set` and `frozenset` types.

- Values of type `set` are mutable: you can add or delete members.
- A `frozenset` value is immutable: you can't change the membership, but you can use a `frozenset` value in contexts where `set` values are not allowed. For example, you can use a `frozenset` as a key in a dictionary, but you can't use a `set` value as a dictionary key.

To create a `set` or `frozenset`, see Section 13.33, “set(): Create an algebraic set” (p. 55) and Section 13.14, “frozenset(): Create a frozen set” (p. 49).

A number of functions that work on sequences also work on sets. In each case, the set is converted to a list before being passed to the function.
• Section 13.2, “\texttt{all()}: Are all the elements of a sequence true?” (p. 45). Predicate to test whether all members of a set are \texttt{True}.
• Section 13.3, “\texttt{any()}: Are any of the members of a sequence true?” (p. 45). Predicate to test whether any member of a set is \texttt{True}.
• Section 13.12, “\texttt{filter()}: Extract qualifying elements from a sequence” (p. 48). Returns a list of the elements that pass through a filtering function.
• Section 13.17, “\texttt{iter()}: Produce an iterator over a sequence” (p. 50). Returns an iterator that will visit every element of the set.
• Section 13.18, “\texttt{len()}: Number of elements” (p. 50). Returns the length (cardinality) of the set.
• Section 13.19, “\texttt{list()}: Convert to a list” (p. 51). Returns the elements of the set as a list.
• Section 13.21, “\texttt{map()}: Apply a function to each element of a sequence” (p. 51). Returns a list containing the result of the application of a function to each element of a set.
• Section 13.22, “\texttt{max()}: Largest element of a sequence” (p. 51). Returns the largest element of a set.
• Section 13.23, “\texttt{min()}: Smallest element of a sequence” (p. 52). Returns the smallest element of a set.
• Section 13.30, “\texttt{reduce()}: Sequence reduction” (p. 54). Returns the result of the application of a given function pairwise to all the elements of a set.
• Section 13.34, “\texttt{sorted()}: Sort a sequence” (p. 56). Returns a list containing the sorted elements of the set.

8.1. Operations on mutable and immutable sets

These operations are supported by both \texttt{set} and \texttt{frozenset} types:

\[ \mathbf{x \ in \ S} \]
Predicate that tests whether element \( x \) is a member of set \( S \).

\[
>> 1 \ \text{in} \ \text{set}([0,1,4]) \\
\text{True} \\
>> 99 \ \text{in} \ \text{set}([0,1,4]) \\
\text{False}
\]

\[ \mathbf{x \ not \ in \ S} \]
Predicate that tests whether element \( x \) is \textit{not} a member of set \( S \).

\[
>> 1 \ \text{not} \ \text{in} \ \text{set}([0,1,4]) \\
\text{False} \\
>> 99 \ \text{not} \ \text{in} \ \text{set}([0,1,4]) \\
\text{True}
\]

\[ \mathbf{S1 \ == \ S2} \]
Predicate that tests whether sets \( S1 \) and \( S2 \) have exactly the same members.

\[
>> \ \text{set}('\texttt{bedac}') \ == \ \text{set}('\texttt{abcde}') \\
\text{True} \\
>> \ \text{set}('\texttt{bedac}') \ == \ \text{set}('\texttt{bedack}') \\
\text{False}
\]

\[ \mathbf{S1 \ != \ S2} \]
Predicate that tests whether sets \( S1 \) and \( S2 \) have different members.

\[
>> \ \text{set}('\texttt{bedac}') \ != \ \text{set}('\texttt{abcde}') \\
\text{False} \\
>> \ \text{set}('\texttt{bedac}') \ != \ \text{set}('\texttt{bedack}') \\
\text{True}
\]
**\( S_1 < S_2 \)**
Predicate that tests whether \( S_1 \) is a proper subset of \( S_2 \); that is, all the elements of \( S_1 \) are also members of \( S_2 \), but there is at least one element of \( S_2 \) that is not in \( S_1 \).

```python
>>> set('ab') < set('ab')
False
>>> set('ab') < set('abcde')
True
```

**\( S_1 > S_2 \)**
Predicate that tests whether \( S_1 \) is a proper superset of \( S_2 \); that is, all the elements of \( S_2 \) are also members of \( S_1 \), but there is at least one element of \( S_1 \) that is not in \( S_2 \).

```python
>>> set('ab') > set('ab')
False
>>> set('abcde') > set('cd')
True
```

**\( S.copy() \)**
Return a new set of the same type as \( S \), containing all the same elements.

```python
>>> s1=set('aeiou')
>>> s2=s1
>>> s3=s1.copy()
>>> s1.add('y')
>>> s1
set(['a', 'e', 'i', 'o', 'u', 'y'])
>>> s2
set(['a', 'e', 'i', 'o', 'u', 'y'])
>>> s3
set(['a', 'i', 'e', 'u', 'o'])
```

**\( S1.difference(S2) \)**
Returns a new set of the same type as \( S1 \), containing only those values found in \( S1 \) but not found in \( S2 \). The \( S2 \) argument may be a set or a sequence.

```python
>>> set('roygbiv').difference('rgb')
set(['i', 'o', 'y', 'v'])
```

**\( S1 - S2 \)**
Same as \( S1.difference(S2) \), except that \( S2 \) must be a set.

```python
>>> set('roygbiv') - set('rgb')
set(['i', 'y', 'v'])
```

**\( S1.intersection(S2) \)**
Returns a new set, of the same type as \( S1 \), containing only the elements found both in \( S1 \) and \( S2 \). \( S2 \) may be a set or a sequence.

```python
>>> set([1,2,3,5,7,11]).intersection(set([1,3,5,7,9]))
set([1, 3, 5, 7])
>>> set([1,3,5]).intersection( (2,4,6,8) )
set([])
```
\textbf{S1 & S2}

Same as \texttt{S1.intersection(S2)}, but \texttt{S2} must be a set.

\textbf{S1.issubset(S2)}

Predicate that tests whether every element of \texttt{S1} is also in \texttt{S2}. \texttt{S2} may be a set or a sequence.

```python
>>> set([1,2]).issubset(set([2,4,1,8]))
True
>>> set([2,4,1,8]).issubset(set([1,2]))
False
>>> set(['r', 'g', 'b']) <= set(['r', 'o', 'y', 'g', 'b', 'i', 'v'])
True
```

\textbf{S1 \leq S2}

Same as \texttt{S1.issubset(S2)}, but \texttt{S2} must be a set.

\textbf{S1.issuperset(S2)}

Predicate that tests whether every element of \texttt{S2} is also in \texttt{S1}. \texttt{S2} may be a set or a sequence.

```python
>>> set([1,2]).issuperset(set([2,4,1,8]))
False
>>> set([2,4,1,8]).issuperset(set([1,2]))
True
```

\textbf{S1 \geq S2}

Same as \texttt{S1.issuperset(S2)}.

\textbf{S1.symmetric_difference(S2)}

Returns a new set of the same type as \texttt{S1}, containing only elements found in \texttt{S1} or \texttt{S2}, but not found in both. The \texttt{S2} argument may be a set or a sequence.

```python
>>> set('aeiou').symmetric_difference('etaoin')
set(['n', 'u', 't'])
```

\textbf{S1 \^ S2}

Same as \texttt{S1.symmetric_difference(S2)}, but \texttt{S2} must be a set.

\textbf{S1.union(S2)}

Returns a new set, with the same type as \texttt{S1}, containing all the elements found in either \texttt{S1} or \texttt{S2}.

The \texttt{S2} argument may be a set or a sequence.

```python
>>> set([1,2]).union(set([1,3,7]))
set([1, 2, 3, 7])
>>> set([1,2]).union( (8,2,4,5) )
set([8, 1, 2, 4, 5])
```

\textbf{S1 | S2}

Same as \texttt{S1.union(S2)}.

\section*{8.2. Operations on mutable sets}

The operations described in this section apply to \texttt{set} (mutable) values, but may not be used with \texttt{frozenset} (immutable) values.

\texttt{S.add(x)}

Add element \texttt{x} to set \texttt{S}. Duplicate elements will be ignored.
```python
In [1]: pbr=set(['USA', 'Brazil', 'Canada'])
In [2]: pbr.add('Australia')
In [3]: pbr
Out[3]: set(['Brazil', 'Canada', 'Australia', 'USA'])
In [4]: pbr.add('USA')
In [5]: pbr
Out[5]: set(['Brazil', 'Canada', 'Australia', 'USA'])
In [6]: pbr.clear()
In [7]: pbr
Out[7]: set([])
In [8]: pbr.discard('Swaziland')
In [9]: pbr
Out[9]: set(['Brazil', 'Australia', 'USA'])
In [10]: pbr.discard('Australia')
In [11]: pbr
Out[11]: set(['Brazil', 'USA'])

S.clear()
Remove all the elements from set S.
```  
```python
In [1]: pbr=set(['Brazil', 'USA'])
In [2]: pbr.clear()
In [3]: pbr
Out[3]: set([])
```

```python
S.discard(x)
If set S contains element x, remove that element from S.
If x is not in S, it is not considered an error; compare S.remove(x).
```  
```python
In [1]: pbr=set(['Brazil', 'Australia', 'USA'])
In [2]: pbr.discard('Swaziland')
In [3]: pbr
Out[3]: set(['Brazil', 'Australia', 'USA'])
In [4]: pbr.discard('Australia')
In [5]: pbr
Out[5]: set(['Brazil', 'USA'])
```  

```python
S1.difference_update(S2)
Modify set S1 by removing any values found in S2. Value S2 may be a set or a sequence.
```  
```python
In [1]: s1=set('roygbiv')
In [2]: s1.difference_update('rgb')
In [3]: s1
Out[3]: set(['i', 'o', 'v', 'y'])
```

```python
S1 -= S2
Same as S1.difference_update(S2), but S2 must be a set.
```  
```python
In [1]: s1=set('roygbiv')
In [2]: s1.intersection_update('roy')
In [3]: s1
Out[3]: set(['y', 'r', 'o'])
```  

```python
S1 & S2
Same as S1.intersection_update(S2), but S2 must be a set.
```
\texttt{S.remove(x)}

If element \(x\) is in set \(S\), remove that element from \(S\).

If \(x\) is not an element of \(S\), this operation will raise a \texttt{KeyError} exception.

```
>>> pbr
set(['Brazil', 'Canada', 'Australia', 'USA'])
>>> pbr.remove('Canada')
>>> pbr
set(['Brazil', 'Australia', 'USA'])
>>> pbr.remove('Swaziland')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
KeyError: 'Swaziland'
```

\texttt{S1.symmetric_difference_update(S2)}

Remove from \(S1\) any elements found in both \(S1\) and \(S2\). Value \(S2\) may be a set or a sequence.

```
>>> s1=set('abcd')
>>> s1.symmetric_difference_update('cdefg')
>>> s1
set(['a', 'b', 'e', 'g', 'f'])
```

\(S1 ^= S2\)

Same as \(S1\).\texttt{symmetric_difference_update(S2)}), but \(S2\) must be a set.

\texttt{S1.update(S2)}

Add to \(S1\) any elements of \(S2\) not found in \(S1\). The \(S2\) argument may be a set or a sequence.

```
>>> s1=set('rgb')
>>> s1
set(['r', 'b', 'g'])
>>> s1.update('roygbiv')
>>> s1
set(['b', 'g', 'i', 'o', 'r', 'v', 'y'])
```

\(S1 |= S2\)

Same as \(S1\).\texttt{update(S2)}, but \(S2\) must be a set.

9. \textbf{Type dict: Dictionaries}

Python dictionaries are one of its more powerful built-in types. They are generally used for look-up tables and many similar applications.

A Python dictionary represents a set of zero or more ordered pairs \((k_i, v_i)\) such that:

- Each \(k_i\) value is called a \textit{key};
- each key is unique and immutable; and
- the associated \textit{value} \(v_i\) can be of any type.

Another term for this structure is \textit{mapping}, since it maps the set of keys onto the set of values (in the algebraic sense).

To create a new dictionary, use this general form:
{ \( k_0 : v_0, k_1 : v_1, \ldots \) }

There can be any number of key-value pairs (including zero). Each key-value has the form “\( k_j : v_j \)”, and pairs are separated by commas. Here are some examples of dictionaries:

\[
\{
\} \\
\{ 'Bolton': 'Notlob', 'Ipswich': 'Esher' \} \\
\{ (1,1):48, (8,20): 52 \}
\]

For efficiency reasons, the order of the pairs in a dictionary is arbitrary: it is essentially an unordered set of ordered pairs. If you display a dictionary, the pairs may be shown in a different order than you used when you created it.

```python
>>> signals = {0:'red', 1: 'yellow', 2:'green'}
>>> signals
{2: 'green', 0: 'red', 1: 'yellow'}
```

### 9.1. Operations on dictionaries

These operations are available on any dictionary object \( D \):

- `len(D)`
  Returns the number of key-value pairs in \( D \).

- `D[k]`
  If dictionary \( D \) has a key whose value is equal to \( k \), this operation returns the corresponding value for that key. If there is no matching key, it raises a `KeyError` exception.

  ```python
  >>> signals = {0:'red', 1: 'yellow', 2:'green'}
  >>> signals[2]
  'green'
  >>> signals[88]
  Traceback (most recent call last):
    File "<stdin>", line 1, in <module>
  KeyError: 88
  ```

- `D[k] = v`
  If dictionary \( D \) does not have a key-value pair whose key equals \( k \), a new pair is added with key \( k \) and value \( v \).
  
  If \( D \) already has a key-value pair whose key equals \( k \), the value of that pair is replaced by \( v \).

- `k in D`
  A predicate that tests whether \( D \) has a key equal to \( k \).

  ```python
  >>> roster={1:'Pat', 2:'Ray', 3:'Min'}
  >>> 3 in roster
  True
  >>> 88 in roster
  False
  ```

- `k not in D`
  A predicate that tests whether \( D \) does not have a key equal to \( k \).

  ```python
  >>> roster={1:'Pat', 2:'Ray', 3:'Min'}
  >>> 3 not in roster
  ```
del \ D[k]

In Python, \texttt{del} is a statement, not a function; see Section 15.3, “The \texttt{del} statement: Delete a name or part of a value” (p. 73).

If dictionary \( D \) has a key-value pair whose key equals \( k \), that key-value pair is deleted from \( D \). If there is no matching key-value pair, the statement will raise a \texttt{KeyError} exception.

\begin{verbatim}
>>> rgb = {'red': '#ff0000', 'green': '#00ff00', 'blue': '#0000ff'}
>>> del rgb['red']
{'blue': '#0000ff', 'green': '#00ff00'}
>>> del rgb['cerise']
Traceback (most recent call last):
  File "stdin", line 1, in <module>
KeyError: 'cerise'
\end{verbatim}

\( \textbf{D.get}(k, x) \)

If dictionary \( D \) has a key equal to \( x \), it returns the corresponding value, that is, it is the same as the expression “\( D[x] \).”

However, if \( D \) has no key-value pair for key \( k \), this method returns the default value \( x \). The second argument is optional; if omitted, and \( D \) has no key equal to \( k \), it returns \texttt{None}.

\begin{verbatim}
>>> roster={1:'Pat', 2:'Ray', 3:'Min'}
>>> roster.get(2)
'Ray'
>>> v = roster.get(8)
>>> print v
None
>>> roster.get(2, 'Not found')
'Ray'
>>> roster.get(8, 'Not found')
'Not found'
\end{verbatim}

\( \textbf{D.has_key}(k) \)

A predicate that returns \texttt{True} if \( D \) has a key \( k \).

\begin{verbatim}
>>> signals = {0: 'red', 1: 'yellow', 2: 'green'}
>>> signals.has_key(1)
True
>>> signals.has_key(88)
False
\end{verbatim}

\( \textbf{D.items()} \)

Returns the contents of dictionary \( D \) as a list of two-element tuples \((k, v)\), in no particular order.

\begin{verbatim}
>>> signals = {0: 'red', 1: 'yellow', 2: 'green'}
>>> signals.items()
[\( (0, \text{'}red\text{'})\), \( (1, \text{'}yellow\text{'})\), \( (2, \text{'}green\text{'})\)]
\end{verbatim}
\textbf{D.\texttt{iteritems()}}

Returns an iterator that generates the values from dictionary \textit{D} as a sequence of two-element tuples (\textit{k}, \textit{v}). See Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88).

```
>>> roster={1:'Pat', 2:'Ray', 3:'Min'}
>>> rosterScan = roster.iteritems()
>>> for n, name in rosterScan:
...     print '%04d: %s' % (n, name)
...  
0001: Pat
0002: Ray
0003: Min
```

\textbf{D.\texttt{iterkeys()}}

Returns an iterator that generates the keys from dictionary \textit{D}. See Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88).

```
>>> roster={1:'Pat', 2:'Ray', 3:'Min'}
>>> nScan = roster.iterkeys()
>>> for n in nScan:
...     print n,
...  
1 2 3
```

\textbf{D.\texttt{itervalues()}}

Returns an iterator that generates the values from dictionary \textit{D}. See Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88).

```
>>> roster={1:'Pat', 2:'Ray', 3:'Min'}
>>> nameScan = roster.itervalues()
>>> for name in nameScan:
...     print name,
...  
Pat Ray Min
```

\textbf{D.\texttt{keys()}}

Returns a list of the key values in dictionary \textit{D}, in no particular order.

```
>>> signals = {0: 'red', 1: 'yellow', 2: 'green'}
>>> signals.keys()
[1, 0, 2]
```

\textbf{D.\texttt{popitem()}}

Returns an arbitrary entry from dictionary \textit{D} as a \texttt{(key, value)} tuple, and also removes that entry. If \textit{D} is empty, raises a \texttt{KeyError} exception.

```
>>> roster={1:'Pat', 2:'Ray', 3:'Min'}
>>> roster.popitem()
(1, 'Pat')
>>> roster
{2: 'Ray', 3: 'Min'}
```

\textbf{D.\texttt{setdefault}(\textit{k}, \textit{x})}

If dictionary \textit{D} has a key equal to \textit{k}, this method returns the corresponding value \textit{D[k]}. 
If $D$ has no key equal to $k$, the method returns the default value $x$. However, unlike the .get() method, it also creates a new key-value pair $(k, x)$ in $D$.

As with the .get() method, the second argument is optional, and defaults to the value None.

**D.values()**

Returns a list of the values from key-value pairs in dictionary $D$, in no particular order. However, if you call both the .items() and .values() methods of a dictionary without changing that dictionary's contents between those calls, Python guarantees that the ordering of the two results will be the same.

```python
>>> signals = {0: 'red', 1: 'yellow', 2: 'green'}
>>> signals.values()
['yellow', 'red', 'green']
>>> signals.keys()
[1, 0, 2]
```

**D.update(D2)**

Merge the contents of dictionary $D_2$ into dictionary $D$. For any key-value pairs that have the same key in both $D$ and $D_2$, the value for that key in $D$ after this operation will be the value from $D_2$, not the value from $D$.

```python
>>> roster={1:'Pat', 2:'Ray', 3:'Min'}
>>> newer={3:'Bev', 4:'Wes'}
>>> roster.update(newer)
>>> roster
>>> newer
{3: 'Bev', 4: 'Wes'}
```

### 10. Type file: Input and output files

To open a file, use this general form:

```python
f = open(name[,mode[,bufsize]])
```

**name**

The path name of the file to be opened, as a string.

**mode**

An optional string specifying what you plan to do with the file. If omitted, you will get read access to the file. In general the value consists of three parts:

- General mode, one of:
  - r: Read access. The file must already exist. You will not be allowed to write to it.
  - w: Write access. If there is no file by this name, a new one will be created.
  - a: Append access. If there is a file by this name, your initial position will be at the end of the file, and you will be allowed to write (and read). If there is no file by this name, a new one will be created.

**Important**

*If there is an existing file, it will be deleted!*
created. On some systems, all writes to a file with append mode are added at the end of the file, regardless of the current file position.

• If you plan to modify the file, append a “+” next.

For example, mode “r+” puts you at the beginning of an existing file and allows you to write to the file anywhere.

Mode “w+” is the same as “w”: it deletes an existing file if there is any, then creates a new file and gives you write access.

Mode “a+” allows you to write new data at the end of an existing file; if no file by this name exists, it will create a new one.

• If you are handling binary data, as opposed to lines of text, add “b” at the end of the mode string.

\textbf{bufsize}

Buffer size: this affects when physical device writes are done, compared to write operations that your program performs.

• In most cases you will probably want to omit this argument. The default is to use line buffering for terminal-type devices, or some system default for other devices.

• Use 0 to force unbuffered operation. This may be inefficient, but any file writes are performed immediately.

• Use 1 for line buffering: output lines are written whenever you write a line terminator such as ‘\n’.

• Use larger values to specify the actual size of the buffer.

If you are reading text files, and you don’t want to worry about the variety of line termination protocols, you may use a mode value of “U” for “universal line terminator mode.” In this mode, input lines may be terminated with either carriage return (‘\r’), newline (‘\n’), or both, but the lines you receive will always be terminated with a single newline. (Exception: If the last line is unterminated, the string you get will also be unterminated.)

There are a number of potential error conditions. For modes starting with “r”, the file must exist before you open it. Also, you must have access according to the underlying operating system. For example, in Linux environments, you must have read access to read a file, and you must have write access to modify or delete a file. These sorts of failures will raise an IOError exception.

A file is its own iterator (see Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88)). Hence, if you have a file inFile opened for reading, you can use a for loop that looks like this to iterate over the lines of the file:

\begin{verbatim}
for line in inFile:
    ...
\end{verbatim}

The variable line will be set to each line of the file in turn. The line terminator character (if any) will be present in that string.

Other aspects of files:

• Every open file has a current position. Initially, this will be 0L if you opened it for reading or writing, or the size of the file if you opened it with append access. Each write or read operation moves this position by the amount read or written. You can also query or set the file position; see Section 10.1, “Methods on file objects” (p. 42).

• Files may use a technique called buffering. Because physical access to some storage media (such as disk drives) takes a relatively long time, Python may employ a storage area called a buffer as a holding area for data being input or output.
For example, if you are writing data to a disk file, Python may keep the data in the file’s buffer area until the buffer is full and only then actually write it to the physical disk. There are various techniques for controlling this behavior; see Section 10.1, “Methods on file objects” (p. 42).

### 10.1. Methods on file objects

Use these methods on an open file instance \( F \).

- **\( F\).close()**
  Close file \( F \). Any unwritten data in the buffer will be flushed. No further operations will be allowed on the file unless it is reopened with the \( \text{open()} \) function.

- **\( F\).flush()**
  For buffered files, you can use this method to make sure that all data written to the file has been physically transmitted to the storage medium. Closing a file will also flush the buffers. Avoid using this method unless you really need it, as it may make your program less efficient.

- **\( F\).isatty()**
  A predicate that tests whether \( F \) is a terminal; “tty” is an ancient term that originally meant “Tele-type”, but has come to mean any terminal or simulated terminal.

- **\( F\).read(\( n \))**
  Read the next \( n \) characters from \( F \) and return them as a string.
  If there are fewer than \( n \) characters remaining after your current position, you will get all remaining characters. If you are at the end of the file, you will get back an empty string (‘’).
  The argument is optional. If omitted, you will get the entire remaining contents of the file as a string.

- **\( F\).readline(\( maxlen \))**
  Read the next text line from \( F \) and return it as a string, including the line terminator if any.
  If you need to limit the maximum size of incoming lines, pass that size limit as the optional \( maxlen \) argument. The default is to return a line of any size (subject to memory limitations). If the line exceeds \( maxlen \), the file position will be left pointing to the first unread character of that line.

- **\( F\).readlines()**
  Read all remaining lines from \( F \) and return them as a list of strings, including line terminator characters if any.

- **\( F\).seek(\( offset \), whence)***
  Change the file’s current position. The value of \( whence \) determines how the \( offset \) value is used to change the position:
  - 0: Set the position to \( offset \) bytes after the beginning of the file.
  - 1: Move the current position \( offset \) bytes toward the end of the file.
  - 2: Move the current position \( offset \) bytes relative to the end of the file.

  For example, for a file \( f \), this operation would set the position to 4 bytes before the end of the file:

  \[
  f\text{.seek}(-4, 2)
  \]

- **\( F\).tell()**
  This method returns the current file position relative to the beginning as a \( \text{long} \) value.

  ```
  >>> f=open('roundtable', 'w')
  >>> f.write('Bedevere')
  >>> f.tell()
  ```
>>> f.seek(2L)
>>> f.tell()
2L

**F.truncate([pos])**
Remove any contents of F past position *pos*, which defaults to the current position.

**F.write(s)**
Write the contents of string *s* to file *F*. This operation will not add terminator characters; if you want newlines in your file, include them in the string *s*.

**F.writelines(S)**
For a sequence *S* containing strings, write all those strings to *F*. No line terminators will be added; you must provide them explicitly if you want them.

# 11. None: The special placeholder value

Python has a unique value called **None**. This special null value can be used as a placeholder, or to signify that some value is unknown.

In conversational mode, any expression that evaluates to **None** is not printed. However, if a value of **None** is converted to a string, the result is the string '**None**'; this may happen, for example, in a `print` statement.

```python
>>> x = None
>>> x
>>> print x
None
```

The value **None** is returned from any function that executes a `return` statement with no value, or any function after it executes its last line if that last line is not a `return` statement.

```python
>>> def useless():
...     print "Useless!"
...
>>> useless()
Useless!
>>> z = useless()
Useless!
>>> z
>>> print z
None
>>> def boatAnchor():
...     pass
...
>>> x = boatAnchor()
>>> x
>>> print x
None
```
12. Operators and expressions

Python’s operators are shown here from highest precedence to lowest, with a ruled line separating groups of operators with equal precedence:

Table 6. Python operator precedence

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>(E)</code></td>
<td>Parenthesized expression or tuple.</td>
</tr>
<tr>
<td><code>[E, ...]</code></td>
<td>List.</td>
</tr>
<tr>
<td><code>{key:value, ...}</code></td>
<td>Dictionary.</td>
</tr>
<tr>
<td><code>\</code> ` `</td>
<td>Convert to string representation.</td>
</tr>
<tr>
<td><code>x.attribute</code></td>
<td>Attribute reference.</td>
</tr>
<tr>
<td><code>x[...]</code></td>
<td>Subscript or slice; see Section 7.1, “Operations common to all the sequence types” (p. 10).</td>
</tr>
<tr>
<td><code>f(...)</code></td>
<td>Call function f.</td>
</tr>
<tr>
<td><code>x**y</code></td>
<td>x to the y power.</td>
</tr>
<tr>
<td><code>-x</code></td>
<td>Negation.</td>
</tr>
<tr>
<td><code>~x</code></td>
<td>Bitwise not (one’s complement).</td>
</tr>
<tr>
<td><code>x*y</code></td>
<td>Multiplication.</td>
</tr>
<tr>
<td><code>x/y, x//y</code></td>
<td>Division. The “//” form discards the fraction from the result. For example, “13.9//5.0” returns the value 2.0.</td>
</tr>
<tr>
<td><code>x%y</code></td>
<td>Modulo (remainder of x/y).</td>
</tr>
<tr>
<td><code>x+y</code></td>
<td>Addition, concatenation.</td>
</tr>
<tr>
<td><code>x-y</code></td>
<td>Subtraction.</td>
</tr>
<tr>
<td><code>x&lt;&lt;y</code></td>
<td>x shifted left y bits.</td>
</tr>
<tr>
<td><code>x&gt;&gt;y</code></td>
<td>x shifted right y bits.</td>
</tr>
<tr>
<td><code>x&amp;y</code></td>
<td>Bitwise and.</td>
</tr>
<tr>
<td><code>x^y</code></td>
<td>Bitwise exclusive or.</td>
</tr>
<tr>
<td>`x</td>
<td>y`</td>
</tr>
<tr>
<td><code>x&lt;y, x&lt;=y, x&gt;y, x&gt;=y, x!=y, x==y</code></td>
<td>Comparisons. These operators are all predicates; see Section 12.1, “What is a predicate?” (p. 44).</td>
</tr>
<tr>
<td><code>x in y, x not in y</code></td>
<td>Test for membership.</td>
</tr>
<tr>
<td><code>x is y, x is not y</code></td>
<td>Test for identity.</td>
</tr>
<tr>
<td><code>not x</code></td>
<td>Boolean “not.”</td>
</tr>
<tr>
<td><code>x and y</code></td>
<td>Boolean “and.”</td>
</tr>
<tr>
<td><code>x or y</code></td>
<td>Boolean “or.”</td>
</tr>
</tbody>
</table>

12.1. What is a predicate?

We use the term predicate to mean any Python function that tests some condition and returns a Boolean value.

For example, `x < y` is a predicate that tests whether `x` is less than `y`. For example, `5 < 500` returns True, while `5 >= 500` returns False.
13. Basic functions

Python has a lot of built-in functions. This section describes the ones that most people use most of the time. If you are interested in exploring some of the more remote corners of Python, see Section 14, “Advanced functions” (p. 59).

13.1. abs() : Absolute value

To find the absolute value of a number \( x \):

\[
\text{abs}(x)
\]

If \( x \) is negative, the function returns \(-x\); otherwise it returns \( x \).

For complex values, the function returns the magnitude, that is, the square root of \((x\cdot\text{real}^2+x\cdot\text{imag}^2)\).

```python
>>> abs(-33)
33
>>> abs(33)
33
>>> abs(0)
0
>>> abs(complex(1,5))
5.0990195135927845
```

13.2. all() : Are all the elements of a sequence true?

A predicate that tests whether all the elements of some sequence are considered True. If any elements are not already type bool, they are converted to Boolean values using the bool() built-in function.

```python
>>> all([True, 14, (88,99)])
True
>>> all([True, 14, (88,99), None])
False
```

13.3. any() : Are any of the members of a sequence true?

This function, applied to some sequence, is a predicate that tests whether any of the elements of that sequence are True. If any element is not already type bool, it is converted to a Boolean value using the bool() built-in function.

```python
>>> noneTrue = (0, 0.0, (), None)
>>> any(noneTrue)
False
>>> someTrue = (0, 0.0, (88,), 'penguin')
>>> any(someTrue)
True
```
13.4. bool(): Convert to Boolean

This function takes any value $x$ and converts it to a Boolean (true or false) value.

For the list of values that are considered True or False, see Section 6.3, “Type bool: Boolean truth values” (p. 8). Examples:

```python
>>> bool(0)
False
>>> bool(0.0)
False
>>> bool(0L)
False
>>> bool(0j)
False
>>> bool('')
False
>>> bool([])
False
>>> bool({})
False
>>> bool(0)
False
>>> bool(1)
True
>>> bool(15.9)
True
>>> bool([0])
True
>>> bool((None,))
True
>>> bool({None: False})
True
```

13.5. chr(): Get the character with a given code

For arguments $n$ in the range 0 to 255, this function returns a one-character string containing the character that has code $n$. Compare Section 13.26, “ord(): Find the numeric code for a character” (p. 52).

```python
>>> chr(65)
'A'
>>> chr(0)
'\x00'
```

13.6. cmp(): Compare two values

This function compares the values of two arguments $x$ and $y$:

```python
cmp(x, y)
```

The return value is:
• A negative number if $x$ is less than $y$.
• Zero if $x$ is equal to $y$.
• A positive number if $x$ is greater than $y$.

The built-in `cmp()` function will typically return only the values -1, 0, or 1. However, there are other places that expect functions with the same calling sequence, and those functions may return other values. It is best to observe only the sign of the result.

```
>>> cmp(2,5)
-1
>>> cmp(5,5)
0
>>> cmp(5,2)
1
>>> cmp('aardvark', 'aardwolf')
-1
```

13.7. `complex()`: Convert to complex type

To create a complex number from a real part $R$ and a complex part $I$:

```
complex($R$, $I$)
```

Example:

```
>>> complex(-4.04, 3.173)
(-4.04+3.173j)
```

13.8. `dict()`: Convert to a dictionary

This function creates a new dictionary from its arguments. The general form is:

```
dict($v$, $k_0$=$v_0$, $k_1$=$v_1$, ...)
```

That is, there may be one optional positional argument or any number of keyword arguments.

• If you supply no arguments, you get a new, empty dictionary.
• If a positional argument is supplied, it must be a sequence of two-element sequences. Each two-element sequence becomes one key-value pair of the result.

```
>>> dict()
{}
>>> dict ( [ (0, 'stop'), (1, 'go') ] )
{0: 'stop', 1: 'go'}
>>> dict(((y', 'boy'), ('x', 'girl')))
{'y': 'boy', 'x': 'girl'}
```

• If you supply any keyword arguments, each keyword becomes a key in the resulting dictionary, and that argument's value becomes the corresponding value of that key-value pair.

```
>>> dict(bricks='sleep', keith='maniac', rj='gumby')
{'bricks': 'sleep', 'keith': 'maniac', 'rj': 'gumby'}
```
13.9. divmod(): Quotient and remainder

Sometimes you want both the quotient and remainder when dividing \( x \) by \( y \). This function returns a tuple \((q, r)\), where \( q \) is the quotient and \( r \) is the remainder.

If either \( x \) or \( y \) is a float, the returned value \( q \) is the whole part of the quotient, and the returned \( r \) is computed as \( x - (r*d) \).

\[
\text{divmod}(x, y)
\]

Examples:

\[
\begin{align*}
\text{>>> divmod(13, 5)} & \quad (2, 3) \\
\text{>>> divmod(1.6, 0.5)} & \quad (3.0, 0.10000000000000009)
\end{align*}
\]

13.10. enumerate(): Step through indices and values of a sequence

Given a sequence \( S \), this function produces an iterator that iterates over the pairs of values \((i, v)\) where each \( i \) is a position in \( S \), and each \( v \) is the corresponding value. For more information on iterators, see Section 172, “Iterators: Values that can produce a sequence of values” (p. 88).

\[
\begin{align*}
\text{>>> L = ['Ministry', 'of', 'Silly', 'Walks']}
\text{>>> for where, what in enumerate(L):}
\text{... print "[%d] %s" % (where, what)}
\end{align*}
\]

\[
\begin{align*}
[0] \text{ Ministry} \\
[1] \text{ of} \\
[2] \text{ Silly} \\
[3] \text{ Walks}
\end{align*}
\]

13.11. file(): Open a file

This function is identical to the open() function; for details, see Section 10, “Type file: Input and output files” (p. 40).

13.12. filter(): Extract qualifying elements from a sequence

This function is useful for removing some of the elements of a sequence. You must provide a filtering function that takes one argument and returns a Boolean value. Here is the calling sequence:

\[
\text{filter}(f, S)
\]

The filtering function \( f \) is the first argument. It is applied to every element of some sequence \( S \). The result is a new sequence of the same type as \( S \), containing only those elements \( x \) of \( S \) for which \( f(x) \) returned True.

\[
\begin{align*}
\text{>>> def isOdd(x):}
\text{... if (x%2) == 1: return True}
\text{... else: return False}
\text{...}
\text{>>> filter(isOdd, [88, 43, 65, -11, 202])}
\end{align*}
\]
[43, 65, -11]
>>> filter(isOdd, (1, 2, 4, 6, 9, 3, 3))
(1, 9, 3, 3)
>>> def isLetter(c):
...     return c.isalpha()
...     return
>>> filter(isLetter, "01234abcdeFGHIJ*(&!^")
'abcdeFGHIJ'

13.13. float(): Convert to float type

Converts a value to type float. The argument must be a number, or a string containing a numeric value in string form (possibly surrounded by whitespace). If the argument is not a valid number, this function will raise a ValueError exception. If no argument is given, it will return 0.0.

>>> float()
0.0
>>> float(17)
17.0
>>> float(' 3.1415  ')
3.1415000000000002
>>> print float('6.0221418e23')
6.0221418e+23
>>> float('142x')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: invalid literal for float(): 142x

13.14. frozenset(): Create a frozen set

This function is used to create a new frozenset value: an immutable set. General forms:

frozenset()
frozenset(S)

The first form creates an empty frozenset. To convert an existing sequence value S to a frozenset, use the second form.

>>> frozenset()
frozenset([])
>>> frozenset('aeiou')
frozenset(['a', 'i', 'e', 'u', 'o'])
>>> frozenset([0, 0, 44, 0, 44, 18])
frozenset([0, 18, 44])

For more information, see Section 8, “Types set and frozenset: Set types” (p. 31).

13.15. hex(): Convert to base 16

Given an integer, this function returns a string displaying that value in hexadecimal (base 16).
13.16. int(): Convert to int type

To convert a number of a different type to int type, or to convert a string of characters that represents a number:

```
int(ns)
```

where `ns` is the value to be converted. If `ns` is a float, the value will be truncated, discarding the fraction.

If you want to convert a character string `s` to an int, and that string is not a decimal number, use this form, where `b` is an integer in the range [2, 36] that specifies the base (radix).

```
int(s, b)
```

Examples:

```
>>> int(43L)
43
>>> int(True)
1
>>> int(False)
0
>>> int(43.89)
43
>>> int("69")
69
>>> int('77', 8)
63
>>> int('7ff', 16)
2047
>>> int('10101', 2)
21
```
13.19. `list()` : Convert to a list

This function creates a new list. The argument `x` may be a sequence or an iterator.

```
>>> list(('Bentham', 'Locke', 'Hobbes'))
['Bentham', 'Locke', 'Hobbes']
>>> list("Bruce")
['B', 'r', 'u', 'c', 'e']
```

13.20. `long()` : Convert to long type

This function works exactly the same way as Section 13.16, “`int()` : Convert to int type” (p. 50), except that it produces a result of type `long`.

```
>>> long(43)
43L
>>> long(43.889)
43L
```

13.21. `map()` : Apply a function to each element of a sequence

Given a function \( f \) that takes one argument and returns a value, this function applies \( f \) to each element of a sequence \( S \), and returns a new list containing the results of each of those function applications.

```
>>> def add100(x):
...     return x+100
...     >>> map(add100, (44,22,66))
[144, 122, 166]
```

13.22. `max()` : Largest element of a sequence

Given a sequence \( S \) containing at least one element, \( \text{max}(S) \) returns the largest element of the sequence.
13.23. min(): Smallest element of a sequence

Given a sequence \( S \) containing at least one element, \( \text{max}(S) \) returns the largest element of the sequence.

```python
>>> max('blimey')
'y'
>>> max([-505, -575, -144, -288])
-144
```

13.24. oct(): Convert to base 8

Given a number \( n \), \( \text{oct}(n) \) returns a string containing an octal (base 8) representation of \( n \). Consistent with Python's convention that any number starting with a zero is considered octal, the result of this function will always have a leading zero.

```python
>>> oct(0)
'0'
>>> oct(127)
'0177'
```

13.25. open(): Open a file

Open a file. For the calling sequence, see Section 10, “Type file: Input and output files” (p. 40).

13.26. ord(): Find the numeric code for a character

Given a string \( s \) containing a single character, \( \text{ord}(s) \) returns that character's numeric code. Compare Section 13.5, “\text{chr}(): Get the character with a given code” (p. 46).

```python
>>> ord('A')
65
>>> ord('\x00')
0
```

13.27. pow(): Exponentiation

There are two ways to compute \( x^y \) in Python. You can write it as “\( x**y \)”. There is also a function that does the same thing:

```python
pow(x, y)
```

For integer arithmetic, the function also has a three-argument form that computes \( x^y \% z \), but more efficiently than if you used that expression:

```python
pow(x, y, z)
```
Examples:

```python
>>> 2**4
16
>>> pow(2,4)
16
>>> pow(2.5, 4.5)
61.763235550163657
>>> (2**9)%3
2
>>> pow(2,9,3)
2
```

### 13.28. range(): Generate an arithmetic progression as a list

This function generates a list containing the values of an arithmetic progression, that is, a sequence of numbers such that the difference between adjacent numbers is always the same. There are three forms:

- **range(n)**
  - Returns the list \([0, 1, 2, \ldots, n-1]\). Note that the result never includes the value \(n\).

- **range(start, stop)**
  - Returns the list \([start, start+1, start+2, \ldots, stop-1]\). The result never includes the \(stop\) value.

- **range(start, stop, step)**
  - Returns the list \([start, start+step, start+2*step, \ldots]\), up to but not including the value of \(step\). The value of \(step\) may be negative.

Examples:

```python
>>> range(4)
[0, 1, 2, 3]
>>> range(4,9)
[4, 5, 6, 7, 8]
>>> range(10,104,10)
[10, 20, 30, 40, 50, 60, 70, 80, 90, 100]
>>> range(5,-1,-1)
[5, 4, 3, 2, 1, 0]
```

### 13.29. raw_input(): Prompt and read a string from the user

To prompt for keyboard entry, use this function:

```python
raw_input(p)
```

The argument \(p\) is a prompt string that is written to standard output. Then a line is read from standard input and returned as a string, without its trailing newline character.

```python
>>> party = raw_input('Party affiliation: ')  
Party affiliation: Slightly silly. 
>>> party  
'Slightly silly.'
```
If the user signals end of input (e.g., with `Control-D` under Unix), the function raises an **EOFError** exception.

### 13.30. reduce(): Sequence reduction

This function is used to apply some two-argument function \( f(x, y) \) pairwise to each pair of the elements of a sequence \( S \) in turn, with an optional initial value \( I \). Here are the general forms:

```
reduce(f, S)
reduce(f, S, I)
```

The result depends on the number of elements in \( S \), and whether the initial value \( I \) is supplied. Let's look first at the case where argument \( I \) is not supplied.

- If \( S \) has only one element, the result is \( S[0] \).
- If \( S \) has two elements, the result is \( f(S[0], S[1]) \).
- If \( S \) has three elements, the result is \( f(f(S[0], S[1]), S[2]) \).
- If \( S \) has four or more elements, \( f \) is applied first to \( S[0] \) and \( S[1] \), then to that result and \( S[2] \), and so on until all elements are reduced to a single value.

If an initial value \( I \) is provided, the result is the same as \( \text{reduce}(f, [I]+\text{list}(S)) \).

Some examples:

```python
>>> def x100y(x,y):
...     return x*100+y
...  
>>> reduce(x100y, [15])
15
>>> reduce(x100y, [1,2])
102
>>> reduce(x100y, [1,2,3])
10203
>>> reduce(x100y, (), 44)
44
>>> reduce(x100y, [1], 2)
201
>>> reduce(x100y, [1,2], 3)
30102
>>> reduce(x100y, [1,2,3], 4)
4010203
```

### 13.31. reversed(): Produce a reverse iterator

This function, applied to a sequence value \( S \), returns an iterator that generates the elements of \( S \) in reverse order (see Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88)).

```python
>>> L=[22,44,88]
>>> backL = reversed(L)
>>> for i in backL:
...    print i,
```

```
22 44 88
```
13.32. round(): Round to the nearest integral value

To find the integral value nearest to some value \( x \), use this function:

\[
\text{round}(x)
\]

The value is returned as a float. In the case that the fractional part of \( x \) is exactly one-half, the returned value is the integer farther from zero. Examples:

```python
>>> round(4.1)
4.0
>>> round(4.9)
5.0
>>> round(4.5)
5.0
>>> round(-4.1)
-4.0
>>> round(-4.9)
-5.0
>>> round(-4.5)
-5.0
```

You can also provide an optional second argument that specifies how many digits to preserve after the decimal.

```python
>>> from math import pi
>>> round(pi)
3.0
>>> print round(pi,1)
3.1
>>> print round(pi,2)
3.14
>>> print round(pi,5)
3.14159
>>> round(pi,30)
3.1415926535897931
```

13.33. set(): Create an algebraic set

Use this function to create a set value. Here are the general forms:

```python
set()
set(S)
```

When called with no arguments, this function returns a new, empty set. To build a set from an existing sequence \( S \), use the second form. Examples:

```python
>>> empty = set()
>>> empty
set([])
```
>>> len(empty)
0
>>> swallows=set(['African', 'European'])
>>> swallows
set(['European', 'African'])
>>> len(swallows)
2
>>> set((0, 0, 0, 58, 0, 0, 58, 17))
set([0, 17, 58])

For more information about sets, see Section 8, “Types set and frozenset: Set types” (p. 31).

13.34. sorted(): Sort a sequence

This function, applied to a sequence S, produces a new list containing the elements of S in ascending order (or some other order you specify).

Here is the general form:

```
sorted(S[, cmp[, key[, reverse]]])
```

The `cmp`, `key`, and `reverse` arguments are optional, and have the same meaning as in the `.sort()` method of the list type (see Section 7.4.1, “Methods on lists” (p. 26)).

```
>>> L = ['geas', 'clue', 'Zoe', 'Ann']
>>> sorted(L)
['Ann', 'Zoe', 'clue', 'geas']
>>> def ignoreCase(x,y):
...     return cmp(x.upper(), y.upper())
...
>>> sorted(L, ignoreCase)
['Ann', 'clue', 'geas', 'Zoe']
>>> L
['geas', 'clue', 'Zoe', 'Ann']
```

In the first example above, 'Zoe' precedes 'clue', because all uppercase letters are considered to be less than all lowercase letters. The second example shows how to sort strings as if they were all uppercase. Note in the last line that the original list L is unchanged.

13.35. str(): Convert to str type

To convert any value x to a string, use this general form:

```
str(x)
```

For example:

```
>>> str(17)
'17'
>>> str({'boy': 'Relmond', 'girl': 'Wirdley'})
"{'boy': 'Relmond', 'girl': 'Wirdley'}"
```
13.36. `sum()`: Total the elements of a sequence

This function, applied to a sequence \( S \), returns the sum of its elements. There are two general forms:

\[
\begin{align*}
\text{sum}(S) \\
\text{sum}(S, I)
\end{align*}
\]

In the second form, the summing process starts with the initial value \( I \). Examples:

```python
>>> L=[1,2,3,4]
>>> sum(L)
10
>>> sum(L,1000)
1010
>>> sum(() , 1000)
1000
```

13.37. `tuple()`: Convert to a tuple

To convert some sequence value \( S \) to a tuple, use this general form:

\[
\text{tuple}(s)
\]

The result will be a new tuple with the same elements as \( S \) in the same order.

To create an empty tuple, omit the argument. Examples:

```python
>>> tuple()
()
>>> tuple ( ['swallow', 'coconut'] )
('swallow', 'coconut')
>>> tuple ( 'shrubbery' )
('s', 'h', 'r', 'u', 'b', 'b', 'e', 'r', 'y')
>>> tuple ( ['singleton'] )
('singleton',)
```

13.38. `type()` : Return a value's type

This function can be applied to any value. It returns a type object corresponding to the type of that value.

For built-in types, the type object is the same as the name of the type: \texttt{int}, \texttt{str}, \texttt{list}, and so on. To test whether a value \( x \) is some type \( T \), you can use the predicate “\texttt{type}(x) \text{ is } T”.

If you display a type object in conversational mode, it will look like “\texttt{<type 'T'>}”. Examples:

```python
>>> type(i)
<type 'int'>
>>> type(i) is int
True
>>> type([2,4,8]) is list
True
```
13.39. unichr(): Convert a numeric code to a Unicode character

Given a number \( n \), this function returns the Unicode character that has code point \( n \). For more on Unicode, see Section 7.3, “Type unicode: Strings of 32-bit characters” (p. 24).

```python
>>> unichr(0)
u'\x00'
>>> unichr(ord('A'))
u'A'
>>> unichr(0x3046)
u'\u3046'
>>> unichr(0xe0047)
u'\U000e0047'
```

13.40. unicode(): Convert to a Unicode string

Use this function to convert a string to type unicode.

```python
>>> unicode('Groucestel')
u'Groucestel'
>>> unicode()  
' ' 
```

13.41. xrange(): Arithmetic progression generator

The `xrange()` function has exactly the same arguments as the `range()` function (see Section 13.28, “range(): Generate an arithmetic progression as a list” (p. 53)).

The difference is that `xrange()` is a generator (see Section 17.3, “Generators: Functions that can produce a sequence of values” (p. 89)), while `range()` actually builds a list for its result. This means you can use `xrange()` in situations where you want to generate a large series of the values from an arithmetic progression, but you don't have enough memory to build that series as a list.

```python
>>> for i in xrange(2000000000):
...    print i,
...    if i > 8:
...        break
...
0 1 2 3 4 5 6 7 8
>>> for i in range(2000000000):
...    print i,
...    if i > 8:
...        break
...
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
MemoryError
```
13.42. **zip()**: Combine multiple sequences

Given two or more sequences $S_0, S_1, S_2, \ldots$ with the same number of elements, you can combine them to produce a list of tuples $(v_0, v_1, \ldots)$ such that each $v_0$ comes from $S_0$, each $v_1$ comes from $S_1$, and so on.

The action of this function is probably best explained by examples:

```python
>>> L1=[1,2,3,4]
>>> L2=['a', 'b', 'c', 'd']
>>> zip(L1, L2)
[(1, 'a'), (2, 'b'), (3, 'c'), (4, 'd')]
>>> L3=[10.0, 20.0, 30.0, 40.0]
>>> zip(L1, L2, L3)
[(1, 'a', 10.0), (2, 'b', 20.0), (3, 'c', 30.0), (4, 'd', 40.0)]
```

14. **Advanced functions**

This section describes Python functions that most people will never need. However, the advanced Python programmer may find some of them quite useful.

14.1. **basestring**: The string base class

The class `basestring` is the parent class for the two string types, `str` and `unicode`. It exists not because you can call it (you can’t), but for type testing. To test whether some value `s` is an instance of either type of string, use this predicate:

```python
isinstance(s, basestring)
```

See Section 14.12, “`isinstance()`: Is a value an instance of some class or type?” (p. 63).

14.2. **callable()**: Is this thing callable?

This predicate tests whether some value `x` can be called as a function.

```python
callable(x)
```

Class names can be called to create an instance of the class. Instances can be called if they define a `__call__()` special method; see Section 19.3.5, “`__call__()`: What to do when someone calls an instance” (p. 105).

```python
>>> def someFunction():
...     pass
... >>> callable(someFunction)
True
>>> callable(len)
True
>>> callable(int)
True
>>> callable(42)
False
```
14.3. **classmethod(): Create a class method**

The purpose of the `classmethod()` function is to convert a method into a class method. For a discussion of the purpose and usage of class methods, see Section 19.5, “Class methods” (p. 109).

There are two ways to declare a class method within a class:

- You can use the function decorator syntax to declare that `classmethod` is a decorator for your method. Precede the method definition with a line reading:

  ```python
  @classmethod
  def methodName(cls, ...):
      method body
  ```

- In some older versions of Python without the decorator syntax, you can still declare a class method by placing a line after the method definition, at the same indentation level as the method’s `def` statement, having this form:

  ```python
  methodName = classmethod(methodName)
  ```

14.4. **delattr(): Delete a named attribute**

Use this function to delete an attribute named `A` of some instance `I`. It does not return a value. Here is the general form:

```python
delattr(I, A)
```

For example, if an instance `seabiscuit` has a `rider` attribute, this statement would delete that attribute:

```python
delattr(seabiscuit, 'rider')
```

If the instance has no such attribute, this function will raise an `AttributeError` exception.

14.5. **dir(): Display a namespace's names**

The purpose of the `dir()` function is to find out what names are defined in a given namespace, and return a list of those names. If called without arguments, it returns a list of the names defined in the local namespace. This function can be very handy for finding out what items are in a module or class.

Certain special names are found in most or all namespaces:

- `__doc__` is present in every namespace. Initially `None`, you can store documentation there.
- `__name__` is the name of the current module (minus the `.py` extension). In the top-level script or in conversational mode, this name is set to the string `__main__`.
- `__builtins__` is a list of the names of all built-in functions and variables.

In this example sequence, we’ll show you what is in the global namespace just after starting up Python. Then we’ll import the `math` module and display its names.

```python
>>> dir()
['__builtins__', '__doc__', '__name__']
>>> x=5; forkTail='Tyrannus'
>>> dir()
['__builtins__', '__doc__', '__name__', 'forkTail', 'x']
>>> print __doc__
```
None

```
>>> print __name__
__main__

>>> import math
>>> print math.__name__
math

>>> dir(math)
['__doc__', '__file__', '__name__', 'acos', 'asin', 'atan', 'atan2', 'ceil',
'cos', 'cosh', 'degrees', 'e', 'exp', 'fabs', 'floor', 'fmod', 'frexp',
'hypot', 'ldexp', 'log', 'log10', 'modf', 'pi', 'pow', 'radians', 'sin', 'sinh',
'sqrt', 'tan', 'tanh']
```

```
>>> print math.__doc__
This module is always available. It provides access to the
mathematical functions defined by the C standard.

>>> print math.log10.__doc__
log10(x) -> the base 10 logarithm of x.

>>> print __builtins__
__builtins__

>>> dir(__builtins__)
```

### 14.6. `eval()`: Evaluate an expression in source form

This function evaluates a Python expression from a string. Example:

```
>>> cent=100
>>> eval('cent**3')
1000000
```
If you want to evaluate the expression using different name environments, refer to the official documentation. For related features, see also Section 14.7, “execfile(): Execute a Python source file” (p. 62) and Section 15.4, “The exec statement: Execute Python source code” (p. 73).

14.7. execfile(): Execute a Python source file

To execute a sequence of Python statements in some file \( F \), use this function:

```python
eexecfile(F)
```

The function returns None. For additional features that allow you to control the environment of the executed statements, see the official documentation.

14.8. getattr(): Retrieve an attribute of a given name

Use this function to retrieve an attribute of an instance \( I \), where the attribute’s name is a string \( s \).

```python
ggetattr(I, s[, default])
```

If \( I \) has no attribute whose name matches \( s \):

- If you supplied the optional \( default \) value, that value is returned.
- If you don’t supply a \( default \) value and there is no such attribute in \( I \), this function will raise an AttributeError exception.

Example:

```python
>>> class C:
...     def __init__(self, flavor):
...         self.flavor = flavor
...
>>> c = C('garlicky')
>>> getattr(c, 'flavor')
'garlicky'
>>> getattr(c, 'aroma', 'bland')
'bland'
>>> getattr(c, 'aroma')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
AttributeError: C instance has no attribute 'aroma'
```

14.9. globals(): Dictionary of global name bindings

This function takes no arguments and returns a dictionary that represents the current global namespace. The keys of this dictionary are globally defined names, and each corresponding value is the value for that name. This example starts from a fresh execution of Python in conversational mode, so the global namespace has only the three special names discussed in Section 14.5, “dir(): Display a namespace’s names” (p. 60).

---

9 http://docs.python.org/lib/built-in-funcs.html
10 http://docs.python.org/lib/built-in-funcs.html
14.10. **hasattr()**: Does a value have an attribute of a given name?

This predicate tests to see if some instance \(I\) has an attribute whose name is given by some string \(s\):

\[
\text{hasattr}(I, s)
\]

If this function returns True, you can be sure that the instance has an attribute named \(s\). However, if it returns False, attempts to access an attribute may still succeed, if the class provides dynamic attributes; see Section 19.3.11, “\_\_getattr\_\_()” (p. 106). Example:

```python
>>> class C:
...     def \_\_init\_\_(self, disc):
...         self.disk = disc
...     
... c=C('five')
>>> hasattr(c, 'disk')
True
>>> hasattr(c, 'disc')
False
>>> hasattr(c, 'jukebox')
False
>>> c.jukebox = 'Nine'
>>> hasattr(c, 'jukebox')
True
```

14.11. **id()**: Unique identifier

This function, given any Python value, returns an integer value that uniquely identifies it. In most implementations, it is the value’s physical memory address.

```python
>>> i = 20
>>> id(i)
137727456
```

14.12. **isinstance()**: Is a value an instance of some class or type?

Use this function to test whether some instance \(I\) is an instance of some class \(C\) (or an instance of any ancestor class from which \(C\) inherits). The general form:

\[
isinstance(I, C)
\]
The second argument may be a class name, a type object, or a tuple of class names and type objects. If a tuple, the function will test the instance against each of the class names or type objects.

```python
>>> class C1:
...    pass
...>>> class C2(C1):
...    pass
...>>> c2=C2()
>>> isinstance(c2,C2)
True
>>> isinstance(c2,C1)
True
>>> isinstance(c2,int)
False
>>> isinstance(1,type(55))
True
>>> isinstance(1, (int, float, long))
True
>>> isinstance('Ni', (int, float, long))
False
```

### 14.13. `issubclass()`: Is a class a subclass of some other class?

To test whether some class `C1` is a subclass of another class `C2`, use this predicate:

```python
issubclass(C1, C2)
```

Examples:

```python
>>> class Polygon:
...    pass
...>>> class Square(Polygon):
...    pass
...>>> issubclass(Square, Polygon)
True
>>> issubclass(Polygon, Square)
False
>>> issubclass(Square, Square)
True
>>> issubclass(unicode, basestring)
True
```

### 14.14. `locals()`: Dictionary of local name bindings

This function returns a dictionary containing the names and values of all variables in the local namespace. An example:

```python
>>> def f(a, b=1):
...    c=2
```
...     print locals()
...

>>> f(5)
{'a': 5, 'c': 2, 'b': 1}

For related functions, see Section 14.5, “dir(): Display a namespace’s names” (p. 60) and Section 14.9, “globals(): Dictionary of global name bindings” (p. 62).

14.15. **property(): Create an access-controlled attribute**

The purpose of this function is to create a *property* of a class. A property looks and acts like an ordinary attribute, except that you provide methods that controls access to the attribute.

There are three kinds of attribute access: read, write, and delete. When you create a property, you can provide any or all of three methods that handle requests to read, write, or delete that attribute.

Here is the general method for adding a property named *p* to a new-style class *C*.

```python
class C(...):
    def R(self):
        ...
        read method...
    def W(self, value):
        ...
        write method...
    def D(self):
        ...
        delete method...
    p = property(R, W, D, doc)
    ...
```

where:

- *R* is a method that takes no arguments and returns the effective attribute value. If omitted, any attempt to read that attribute will raise `AttributeError`.

- *W* is a method that takes one argument and sets the attribute to that argument’s value. If omitted, any attempt to write that attribute will raise `AttributeError`.

- *D* is a method that deletes the attribute. If omitted, any attempt to delete that attribute will raise `AttributeError`.

- *doc* is a documentation string that describes the attribute. If omitted, defaults to the documentation string of the *R* method if any, otherwise `None`.

To retrieve a property’s documentation, use this form:

```python
C.p.__doc__
```

where *C* is the class name and *p* is the property name.

As an example, here is a small class that defines a property named *x*:

```python
class C(object):
    def __init__(self):
        self.__x=None
    def getx(self):
        print "+++ getx()"
        return self.__x
    def setx(self, v):
        print "+++ setx()"
        self.__x=v
```

print "+++ setx(%s)" % v
self.__x = v
def delx(self):
    print "+++ delx()"
del self.__x
x=property(getx, setx, delx, "Me property 'x'.")

Assuming that class is defined, here is a conversational example.

```python
>>> c=C()
>>> print c.x
+++ getx()
None
>>> print C.x.__doc__
Me property 'x'.
>>> c.x=15
+++ setx(15)
>>> c.x
+++ getx()
15
>>> del c.x
+++ delx()
>>> c.x
+++ getx()
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  File "<stdin>", line 6, in getx
AttributeError: 'C' object has no attribute '_C__x'
```

### 14.16. `reload()`: Reload a module

This function reloads a previously loaded module (assuming you loaded it with the syntax “`import moduleName`”). It is intended for conversational use, where you have edited the source file for a module and want to test it without leaving Python and starting it again. General form:

```
reload(moduleName)
```

The `moduleName` is the actual name of the module, not a string containing its name. For example, if you have imported a module like this:

```
import parrot
```

you would say “`reload(parrot)`”, not “`reload('parrot')`”.

### 14.17. `repr()`: Representation

The `repr` function attempts to convert any value to a string. Unlike the `str()` function, it attempts to display the value in Python source form, that is, in a form suitable for passing to `eval()` (see Section 14.6, “`eval()`: Evaluate an expression in source form” (p. 61)). It works the same as the `...` operator.

Examples:

```
>>> s='Wensleydale'
>>> print s
```
14.18. `setattr()`: Set an attribute

This function is similar to Section 14.8, “`getattr()`: Retrieve an attribute of a given name” (p. 62): it sets the value of some attribute whose name is `A` from an instance `I` to a new value `V`:

```
sattr(I, A, V)
```

Example:

```python
>>> class C5:
...     def __init__(self, x):
...         self.x = x
...     ...
>>> c=C5(14)
>>> c.x
14
>>> setattr(c, 'x', 19)
>>> c.x
19
```

14.19. `slice()`: Create a slice instance

This function creates an instance of type `slice`. A slice instance `I` can be used to index a sequence `S` in an expression of the form `S[I]`. Here is the general form:

```
slice(start, limit, step)
```

The result is a slice that is equivalent to `start:limit:step`. Use `None` to get the default value for any of the three arguments.

Examples:

```python
>>> r = range(9)
>>> r
[0, 1, 2, 3, 4, 5, 6, 7, 8]
>>> r[::2]
[0, 2, 4, 6, 8]
>>> r[slice(None, None, 2)]
[0, 2, 4, 6, 8]
>>> r[3:7]
[3, 4, 5, 6]
>>> r[slice(3,7)]
[3, 4, 5, 6]
>>> r[1::2]
[1, 3, 5, 7]
```
14.20. `staticmethod()`: Create a static method

The purpose of the `staticmethod` function is to convert a method into a static method. See Section 19.4, “Static methods” (p. 108) for definitions and usage.

Typically you will declare a static method using the decorator syntax, like this:

```python
@staticmethod
def methodName(...):
    method body
```

An alternative is to place a line like this after the method’s definition (at the same indentation level as its `def`):

```python
methodName = staticmethod(methodName)
```

14.21. `super()`: Superclass

The purpose of this function is to retrieve the superclass of a given type or object. Superclasses are beyond the scope of this document; see the online documentation for built-in functions\(^\text{11}\) (scroll down).

14.22. `vars()`: Local variables

This function returns a dictionary that represents a symbol table: its keys are variable names, and each key’s corresponding value is its bound value. The official documentation warns you not to change anything in this dictionary, or Bad Things Will Happen.

There are two forms of call:

- `vars()`: Returns the local symbol table.
- `vars(ns)`: Returns the symbol table of a namespace `ns`, where a namespace can be a module, an instance, or a class.

Compare the similar functions Section 14.5, “`dir()`: Display a namespace’s names” (p. 60), Section 14.14, “`locals()`: Dictionary of local name bindings” (p. 64), and Section 14.9, “`globals()`: Dictionary of global name bindings” (p. 62).

15. Simple statements

Python statement types are divided into two groups. Simple statements, that are executed sequentially and do not affect the flow of control, are described first. Compound statements, which may affect the sequence of execution, are discussed in Section 16, “Compound statements” (p. 76).

\(^{11}\) [http://docs.python.org/lib/built-in-funcs.html](http://docs.python.org/lib/built-in-funcs.html)
Here, for your convenience, is a table of all the Python statement types, and the sections where they are described. The first one, the assignment statement, does not have an initial keyword: an assignment statement is a statement of the form “variable = expression”.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Section 15.1, “The assignment statement: name = expression” (p. 69).</th>
</tr>
</thead>
<tbody>
<tr>
<td>assert</td>
<td>Section 15.2, “The assert statement: Verify preconditions” (p. 72).</td>
</tr>
<tr>
<td>break</td>
<td>Section 16.2, “The break statement: Exit a for or while loop” (p. 78).</td>
</tr>
<tr>
<td>continue</td>
<td>Section 16.3, “The continue statement: Jump to the next cycle of a for or while” (p. 78).</td>
</tr>
<tr>
<td>del</td>
<td>Section 15.3, “The del statement: Delete a name or part of a value” (p. 73).</td>
</tr>
<tr>
<td>elif</td>
<td>Section 16.5, “The if statement: Conditional execution” (p. 80) and Section 16.8, “The try statement: Anticipate exceptions” (p. 82).</td>
</tr>
<tr>
<td>else</td>
<td>Section 16.5, “The if statement: Conditional execution” (p. 80).</td>
</tr>
<tr>
<td>except</td>
<td>Section 16.8, “The try statement: Anticipate exceptions” (p. 82).</td>
</tr>
<tr>
<td>exec</td>
<td>Section 15.4, “The exec statement: Execute Python source code” (p. 73).</td>
</tr>
<tr>
<td>finally</td>
<td>Section 16.8, “The try statement: Anticipate exceptions” (p. 82).</td>
</tr>
<tr>
<td>for</td>
<td>Section 16.4, “The for statement: Iteration over a sequence” (p. 79).</td>
</tr>
<tr>
<td>from</td>
<td>Section 15.6, “The import statement: Use a module” (p. 75).</td>
</tr>
<tr>
<td>global</td>
<td>Section 15.5, “The global statement: Declare access to a global name” (p. 73).</td>
</tr>
<tr>
<td>if</td>
<td>Section 16.5, “The if statement: Conditional execution” (p. 80).</td>
</tr>
<tr>
<td>import</td>
<td>Section 15.6, “The import statement: Use a module” (p. 75).</td>
</tr>
<tr>
<td>pass</td>
<td>Section 15.7, “The pass statement: Do nothing” (p. 76).</td>
</tr>
<tr>
<td>print</td>
<td>Section 15.8, “The print statement: Display output values” (p. 76).</td>
</tr>
<tr>
<td>raise</td>
<td>Section 16.6, “The raise statement: Cause an exception” (p. 81).</td>
</tr>
<tr>
<td>return</td>
<td>Section 16.7, “The return statement: Exit a function or method” (p. 82).</td>
</tr>
<tr>
<td>try</td>
<td>Section 16.8, “The try statement: Anticipate exceptions” (p. 82).</td>
</tr>
<tr>
<td>yield</td>
<td>Section 16.9, “The yield statement: Generate one result from a generator” (p. 85).</td>
</tr>
</tbody>
</table>

### 15.1. The assignment statement: name = expression

The purpose of Python’s assignment statement is to associate names with values in your program. It is the only statement that does not start with a keyword. An assignment statement is a line containing at least one single equal sign (=) that is not inside parentheses.

Here is the general form of an assignment statement:

```
target_0 = target_1 = ... = expression
```

In most cases, there will be one target that is a name. Python will evaluate the expression, reducing it to a single value, and then bind that name to the that value.

A binding is an association between a name and a value. It is important to note that in Python, unlike many other languages, names themselves are not associated with a specific type. A name is just a label, and it can be bound to any value of any type at any time. In this example, name x is bound first to an int value 5, then to a str value ‘Some string’.
If a target name was already bound to a value, the name is unbound from that value before it is rebound to the new value. For each value in a running program, Python keeps track of how many names are bound to that value. When the value has no more names bound to it, the value’s memory is automatically recycled. If the value is an instance of a class, its destructor may be called; see Section 19.3.8, \"__del__ ()\": Destructor” (p. 106).

There are several other forms of assignment statement.

\[ n_0 = n_1 = \ldots = \text{expression} \]

If you supply multiple target names, each target will be assigned the value of the \text{expression}. Example:

```python
>>> i = j = errorCount = 0
>>> i
0
>>> j
0
>>> errorCount
0
```

\[ n_0, n_1, \ldots = \text{expression} \]

If the target is a comma-separated list of names, the \text{expression} must evaluate to a sequence with the same number of values as there are names. Each value in the \text{expression} is then bound to the corresponding name. Example:

```python
>>> L = ["Bell's Blue", "male", 11]
>>> name, sex, age = L
>>> name
"Bell's Blue"
>>> sex
'male'
>>> age
11
```

This feature, called “unpacking,” generalizes to arbitrarily nested sequences within sequences. You may group targets inside parentheses (\ldots) or brackets [\ldots] to show the levels of nesting. Here is an example:

```python
>>> s = [1, [2, 3, [4, 5], 6], 7]
>>> a, (b, c, [d, e], f), g = s
>>> print a,b,c,d,e,f,g
1 2 3 4 5 6 7
```

All the assignments are effectively simultaneous. Therefore, you can safely exchange the values of two variables using a statement like this:

```python
v_1, v_2 = v_2, v_1
```

Examples:
name[i] = expression
If name is a list (or other mutable sequence), the expression i must evaluate to an integer. The element at position i is replaced by the value of the expression.

>>> L = range(6)
>>> L
[0, 1, 2, 3, 4, 5]
>>> L[2]
2
>>> L[2] = 888
>>> L
[0, 1, 888, 3, 4, 5]

As the last two lines show, if the dictionary already has a key-value pair for key i, the old value of that pair is replaced by the expression value.

name[start:end] = S
If name is a list or other mutable sequence, you can replace the elements of a slice of that sequence with the elements from some sequence S. (For an explanation of slicing, see Section 7.1, “Operations common to all the sequence types” (p. 10).) This may result in addition, deletion, or replacement of the elements of name. Some examples will give the flavor of this kind of assignment.

>>> L=range(6)
>>> L
[0, 1, 2, 3, 4, 5]
>>> L[2:4]
[2, 3]
>>> L[2:4] = [111, 222, 333, 444, 555]
>>> L
[0, 1, 111, 222, 333, 444, 555, 4, 5]
>>> L[3]
Note
The “=” signs in an assignment is not an operator, as it is in some other languages. You cannot assign a value to a name inside an expression; an assignment statement must stand alone.

```python
>>> a = 5 + (a=7)
File "<stdin>", line 1
  a = 5 + (a=7)
^ SyntaxError: invalid syntax
```

Python also supports augmented assignment. In this form, you may place certain operators before the “=”.
Here is the general form:

```python
name operator= expression
```

An assignment of this general form has the same semantics as this form:

```python
name = name operator expression
```

Supported operator symbols include:

```python
+ - * / % ** >> << & ^ |
```

Examples:

```python
>>> i = 1
>>> i += 3
>>> i
4
>>> i *= 5
>>> i
20
```

15.2. The assert statement: Verify preconditions

To check for “shouldn’t happen” errors, you can use an assert statement:

```python
assert e1
assert e1, e2
```
where $e_1$ is some condition that should be true. If the condition is false, Python raises an Assertion-Error exception (see Section 18, “Exceptions: Error signaling and handling” (p. 90)).

If a second expression $e_2$ is provided, the value of that expression is passed with the exception.

Assertion checking can be disabled by running Python with the -O (optimize) option.

15.3. The del statement: Delete a name or part of a value

The purpose of the del statement is to delete things. The general form is:

```
del L_0, L_1, ...
```

where each $L_i$ is an item to be deleted. You can delete:

- A name. For example, the statement
  ```
del i, j
```
  causes names i and j to become unbound, that is, undefined.

- An element or slice from a list. For example:
  ```
del L[5], M[-2:]
```
  would delete the sixth element of list L and the last two elements of list M.

- One entry in a dictionary. For example, if D is a dictionary,
  ```
del D['color']
```
  would delete from D the entry for key 'color'.

15.4. The exec statement: Execute Python source code

To dynamically execute Python code, use a statement of this form:

```
exec $E_0$ [in $E_1$ [, $E_2$]]
```

Expression $E_0$ specifies what to execute, and may be a string containing Python source code, an open file, or a code object. If $E_1$ is omitted, the code is executed in the local scope. If $E_1$ is given but $E_2$ is not, $E_1$ is a dictionary used to define the names in the global and local scopes. If $E_2$ is given, $E_1$ is a dictionary defining the global scope, and $E_2$ is a dictionary defining the local scope.

15.5. The global statement: Declare access to a global name

The purpose of the global statement is to declare that a function or method intends to change the value of a name from the global scope, that is, a name from outside the function.

When Python reads the definition of a function, it checks each name to see if that name's value may possibly be changed anywhere in the function—that is, if the name shows up on the left side of an assignment statement, or as the induction variable in a for loop, or in any other context where the name’s value can be changed.

Such names are assumed to be local to the function unless you override this behavior by declaring that name in a global statement. Here is the general form:
Some conversational examples may help make this clear. Suppose you define a global variable \( x \); you can use that name inside a function.

```python
>>> x = 5
>>> def show1(): ...
...   print x
... >>> show1()
5
```

However, if you assign a value to \( x \) inside the function, the name \( x \) is now local to the function. It is said to *shadow* the global variable with the same name, and any changes to the value associated with that name inside the function will operate on a local copy, and will not affect the value of the global variable \( x \).

```python
>>> x = 5
>>> def show2(): ...
...   x = 42
...   print x
... >>> show2()
42
>>> x
5
```

But if you actually do want to change the value of the global variable inside the function, just declare it global like this:

```python
>>> x = 5
>>> def show3(): ...
...   global x
...   x = 42
...   print x
... >>> show3()
42
>>> x
42
```

Notice what happens in this case:

```python
>>> x = 5
>>> def show4(): ...
...   print x, "Before"
...   x = 42
...   print x, "After"
... >>> show4()
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  File "<stdin>", line 2, in show4
UnboundLocalError: local variable 'x' referenced before assignment
```
Because the line “x = 42” changes the value of x, and because it is not declared as a global, execution fails because the value of the local variable x is used before that variable has had a value assigned to it.

### 15.6. The import statement: Use a module

One of the cornerstones of Python’s design philosophy is to keep the language relatively small and well-defined, and move all non-essential functionality to library modules. The `import` and `from` statements allow your programs to use items from these library modules.

- Your Python installation will come with a large collection of released modules.
- You can also create your own modules. Just place Python statements defining variables, functions, and classes into a file whose name ends in “.py”.

There are two different statements you can use to import items from a module:

- The `from` statement copies items from a module into your namespace. After importing an item in this way, you can refer to the item simply by its name.

  General forms:

  ```python
  from moduleName import *
  from moduleName import name_1, name_2, ...
  ```

  The first form imports all the items from the module named `moduleName`. If you want to import only specific items, use the second form, and enumerate the names you want from that module.

- The `import` statement makes an entire module’s content available to you as a separate namespace. To refer to some item named `N` in a module named `M`, use the dot notation, `M.N`.

  Here is the general form:

  ```python
  import moduleName, ...
  ```

  If you want to use some module `M` in this way, but you want to change the name to some different name `A`, use this form:

  ```python
  import M as A
  ```

  Here are some examples that use the standard `math` module that is always available in a proper Python install. This module has functions such as `sqrt()` (square root), as well as variables such as `pi`. (Although π is a constant in the mathematical sense, the name `pi` is a variable in the Python sense.)

  ```python
  >>> from math import *
  >>> sqrt(16)
  4.0
  >>> pi
  3.1415926535897931
  ```

  If you wanted only the `sqrt` function and the variable `pi`, this statement would do the job:

  ```python
  from math import sqrt, pi
  ```

  Now some examples of the second form.

  ```python
  >>> import math
  >>> math.sqrt(25)
  5.0
  ```
>>> math.pi
3.1415926535897931

Suppose your program already used the name math for something else, but you still want to use functions from the math module. You can import it under a different name like this:

>>> import math as crunch
>>> crunch.sqrt(25)
5.0
>>> crunch.pi
3.1415926535897931

15.7. The pass statement: Do nothing

Python’s pass statement is a placeholder. It does nothing.

Here’s an example. Suppose you have a function named arr() that does something and then returns a True or False value. You want to keep calling this function until it returns a false value. This code would suffice:

```python
while arr():
    pass
```

15.8. The print statement: Display output values

Use this statement to display values on the standard output stream. General form:

```python
print thing1, thing2, ..., [,]
```

Each thing must be a string, or a value that can be converted into a string by the str() function (see Section 13.35, “str(): Convert to str type” (p. 56)). These strings are written to the standard output stream, with one space between each value. A print statement by itself prints an empty line.

```python
>>> print 4.3, 'Sir Robin', 1./7
4.3 Sir Robin 0.142857142857
>>> for i in range(4):
...    print i**4,
...        ...
0 1 16 81
```  

Normally, a newline is printed after the last value. However, you can suppress this behavior by appending a comma to the end of the list. For example, this statement:

```python
print 'State your name:',
```

would print the string followed by one space and leave the cursor at the end of that line.

16. Compound statements

The statements in this section alter the normal sequential execution of a program. They can cause a statement to be executed only under certain circumstances, or execute it repeatedly.
16.1. Python's block structure

One unusual feature of Python is the way that the indentation of your source program organizes it into blocks within blocks within blocks. This is contrary to the way languages like C and Perl organize code blocks by enclosing them in delimiters such as braces `{ ... }`.

Various Python branching statements like `if` and `for` control the execution of blocks of lines.

- At the very top level of your program, all statements must be unindented—they must start in column one.
- Various Python branching statements like `if` and `for` control the execution of one or more subsidiary blocks of lines.
- A block is defined as a group of adjacent lines that are indented the same amount, but indented further than the controlling line. The amount of indentation of a block is not critical.
- You can use either spaces or `tab` characters for indentation. However, mixing the two is perverse and can make your program hard to maintain. Tab stops are assumed to be every eight columns.

Blocks within blocks are simply indented further. Here is an example of some nested blocks:

```python
if i < 0:
    print "i is negative"
else:
    print "i is nonnegative"
    if i < 10:
        print "i has one digit"
    else:
        print "i has multiple digits"
```

If you prefer a more horizontal style, you can always place statements after the colon (`:`) of a compound statement, and you can place multiple statements on a line by separating them with semicolons (`;`).

Example:

```python
>>> if 2 > 1: print "Math still works"; print "Yay!"
... else: print "Huh?"
...
Math still works
Yay!
```

You can't mix the block style with the horizontal style: the consequence of an `if` or `else` must either be on the same line or in a block, never both.

```python
>>> if 1: print "True"
... print "Unexpected indent error here."
File "stdin", line 2
    print "Unexpected indent error here."
^  
IndentationError: unexpected indent
>>> 
```
16.2. The break statement: Exit a for or while loop

The purpose of this statement is to jump out of a for or while loop before the loop would terminate otherwise. Control is transferred to the statement after the last line of the loop. The statement looks like this:

```
break
```

Here's an example.

```python
>>> for i in [1, 71, 13, 2, 81, 15]:
... print i,
... if (i%2) == 0:
... break
...
1 71 13 2
```

Normally this loop would be executed six times, once for each value in the list, but the break statement gets executed when `i` is set to an even value.

16.3. The continue statement: Jump to the next cycle of a for or while

Use a continue statement inside a for or while loop when you want to jump directly back to the top of the loop and go around again.

- If used inside a while loop, the loop's condition expression is evaluated again. If the condition is False, the loop is terminated; if the condition is True, the loop is executed again.
- Inside a for loop, a continue statement goes back to the top of the loop. If there are any values remaining in the sequence of values that controls the loop, the loop variable is set to the next value in sequence, and the loop body is entered.

If the continue is executed during the last pass through the loop, control goes to the statement after the end of the loop.

Examples:

```python
>>> i = 0
>>> while i < 10:
... print i,
... i += 1
... if (i%3) != 0:
... continue
... print "num",
...
0 1 2 num 3 4 5 num 6 7 8 num 9
>>> for i in range(10):
... print i,
... if (i%4) != 0:
... continue
... print "whee",
...
0 whee 1 2 3 4 whee 5 6 7 8 whee 9
```
16.4. The for statement: Iteration over a sequence

Use a for statement to execute a block of statements repeatedly. Here is the general form. (For the definition of a block, see Section 16.1, “Python’s block structure” (p. 77.).)

```python
for V in S:
    B
```

- V is a variable called the induction variable.
- S is any expression that produces a value that is a sequence, including iterators (see Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88)) and generators (see Section 17.3, “Generators: Functions that can produce a sequence of values” (p. 89)).

    This sequence is called the controlling sequence of the loop.

- B is a block of statements.

The block is executed once for each value in S. During each execution of the block, V is set to the corresponding value of S in turn. Example:

```python
>>> for color in ['black', 'blue', 'transparent']:
...     print color
...
black
blue
transparent
```

In general, you can use any number of induction variables. In this case, the members of the controlling sequence must themselves be sequences, which are unpacked into the induction variables in the same way as sequence unpacking as described in Section 15.1, “The assignment statement: name = expression” (p. 69). Here is an example.

```python
>>> fourDays = ( ('First', 1, 'orangutan librarian'),
...     ('Second', 5, 'loaves of dwarf bread'),
...     ('Third', 3, 'dried frog pills'),
...     ('Fourth', 2, 'sentient luggages') )
>>> for day, number, item in fourDays:
...     print ( "On the %s day of Hogswatch, my true love gave to me" %
...             day + " ", "%d %s"
...             % (number, item) )
...
On the First day of Hogswatch, my true love gave to me 1 orangutan librarian
On the Second day of Hogswatch, my true love gave to me 5 loaves of dwarf bread
On the Third day of Hogswatch, my true love gave to me 3 dried frog pills
On the Fourth day of Hogswatch, my true love gave to me 2 sentient luggages
```

For those of you who like to break things, be advised that you can change the induction variable inside the loop, but during the next pass through the loop, it will be set to the next element of the controlling sequence normally. Modifying the control sequence itself won’t change anything; Python makes a copy of the control sequence before starting the loop.
>>> for i in range(4):
...     print "Before:", i,
...     i += 1000
...     print "After:", i
... Before: 0 After: 1000
Before: 1 After: 1001
Before: 2 After: 1002
Before: 3 After: 1003
>>> L = [7, 6, 1912]
>>> for n in L:
...     L = [44, 55]
...     print n
... 7
6
1912

16.5. The if statement: Conditional execution

The purpose of the if construct is to execute some statements only when certain conditions are true.

Here is the most general form of an if construct:

```python
if E_0:
    B_0
elif E_1:
    B_1
elif \ldots:
    \ldots
else:
    B_f
```

In words, this construct means:

- If expression $E_0$ is true, execute block $B_0$.
- If expression $E_0$ is false but $E_1$ is true, execute block $B_1$.
- If there are more `elif` clauses, evaluate each one’s expression until that expression has a true value, and then execute the corresponding block.
- If all the expressions in `if` and `elif` clauses are false, execute block $B_f$.

An if construct may have zero or more `elif` clauses. The `else` clause is also optional.

Examples:

```python
>>> for i in range(5):
...     if i == 0:
...         print "i is zero",
...     elif i == 1:
...         print "it's one",
...     elif i == 2:
...         print "it's two",
...     else:
...         print "many",
```
16.6. **The raise statement: Cause an exception**

Python’s exception mechanism is the universal framework for dealing with errors—situations where your program can’t really proceed normally. For an overview, see Section 18, “Exceptions: Error signaling and handling” (p. 90).

There are three forms of the `raise` statement:

```
raise
raise E1
raise E1, E2
```

The first form is equivalent to “`raise None, None`” and the second form is equivalent to “`raise E1, None`”. Each form raises an exception of a given type and with a given value. The type and value depend on how many expressions you provide:

<table>
<thead>
<tr>
<th><strong>E1</strong></th>
<th><strong>E2</strong></th>
<th>Exception type</th>
<th>Exception value</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>None</td>
<td>Re-raise the current exception, if any. This might be done, for example, inside an <code>except</code>, <code>else</code>, or <code>finally</code> block; see Section 16.8, “The <code>try</code> statement: Anticipate exceptions” (p. 82).</td>
<td></td>
</tr>
<tr>
<td>class</td>
<td>None</td>
<td><code>E1</code></td>
<td><code>E1()</code></td>
</tr>
<tr>
<td>class</td>
<td>instance of <code>E1</code></td>
<td><code>E1</code></td>
<td><code>E2</code></td>
</tr>
<tr>
<td>class</td>
<td><code>tuple</code></td>
<td><code>E1</code></td>
<td><code>E1(*E2)</code></td>
</tr>
<tr>
<td>class</td>
<td>none of the above</td>
<td><code>E1</code></td>
<td><code>E1(E2)</code></td>
</tr>
<tr>
<td>instance</td>
<td>None</td>
<td>type(<code>E1</code>)</td>
<td><code>E1</code></td>
</tr>
</tbody>
</table>

The current recommended practice is to use a `raise` statement of this form:

```
raise E(message)
```

where `E` is some class derived from the built-in `Exception` class: you can use one of the built-in exceptions, or you can create your own exception classes.
For classes derived from `Exception`, the constructor takes one argument, an error message—that is, a string explaining why the exception was raised. The resulting instance makes that message available as an attribute named `.message`. Example:

```python
>>> try:
...   raise ValueError('The day is too frabjous.')
... except ValueError, detail:
...   pass
...
>>> type(detail)
<type 'exceptions.ValueError'>
>>> detail.message
'The day is too frabjous.'
```

To create your own exceptions, write a class that inherits from `Exception` and passes its argument to the parent constructor, as in this example.

```python
>>> class FashionError(Exception):
...   def __init__(self, sin):
...     Exception.__init__(self, sin)
...
>>> try:
...   print "Let's try shirt 14 and trouser 92."
...   raise FashionError('Clashing plaids.')
...   print "Hey, we got away with it!"
... except FashionError, problem:
...   print "Terminal fashion sin: ", problem.message
...
Let's try shirt 14 and trouser 92.
Terminal fashion sin: Clashing plaids.
```

### 16.7. The return statement: Exit a function or method

Within any function or method, you can exit the function with a `return` statement. There are two forms:

```
return expression
```

In the first form, execution resumes at the point where the function or method was called, and the value of the `expression` is substituted into the calling statement.

The second form is the equivalent of “`return None`”. (See Section 11, “None: The special placeholder value” (p. 43).)

### 16.8. The try statement: Anticipate exceptions

The purpose of a “try” construct is to specify what actions to take in the event of errors. For an introduction to Python’s exception mechanism, see Section 18, “Exceptions: Error signaling and handling” (p. 90).

When an exception is raised, two items are associated with it:

- An exception type, and
- an exception value.
Here is the most general form of a try construct. Symbols like $B_0$ and $B_1$ represent indented blocks of statements. Each except clause names some exception class $E_i$, and optionally a variable $v_i$ that will be set to the exception value.

```
try:
    $B_0$
except $E_1$, $v_1$:
    $B_1$
except $E_2$, $v_2$:
    $B_2$
except ...:
    ...
else:
    $B_e$
finally:
    $B_f$
```

The else: and finally: blocks are optional. There must be at least one except block, but there may be any number.

Here is a simplified description of the execution of a try block in general:

1. If $B_0$ executes without raising any exceptions, the else block $B_e$ is executed, then the finally block $B_f$.
2. If the execution of block $B_0$ raises some exception with type $E_0$, that type is compared to all the exception classes $E_1, E_2, ...$ named in the various except clauses.
   
   The criterion for matching is that $E_0$ must be the same class as, or some subclass of, one of the class names $E_i$ mentioned in an except clause. If there are multiple matches, the first except clause block $B_i$ is executed. In this case, we say that the exception has been caught (or handled).
   
   When an except block matches and is entered, the corresponding variable $v_i$ is set to the value of the exception.
3. If there is a finally clause, it is executed, whether the exception was caught or not. If the exception was not caught, it is re-raised after the end of the finally clause.

Examples:

```
>>> try:
...     raise ValueError, "Your arm's off."
... except IOError, detail:
...     print "I/O error caught: ", detail
... except ValueError, detail:
...     print "Value error caught: ", detail
... except SyntaxError, detail:
...     print "Syntax error caught: ", detail
... else:
...     print "This is the else clause"
... finally:
...     print "This is the finally clause"
...
Value error caught: Your arm's off.
This is the finally clause
>>> try:
...     raise ValueError, "Uncaught!"
```
... except IOError, detail:
...     print "I/O error:", detail
... else:
...     print "This is the else clause"
... finally:
...     print "This is the finally clause"
...
This is the finally clause
Traceback (most recent call last):
  File "<stdin>", line 2, in <module>
ValueError: Uncaught!
>>> try:
...     print "No exceptions raised"
... except ValueError, x:
...     print "ValueError:", x
... else:
...     print "This is the else clause"
... finally:
...     print "This is the finally clause"
...
No exceptions raised
This is the else clause
This is the finally clause

For those of you who are interested in the gory details, the fun begins when a second or even a third exception is raised inside an except, else, or finally clause. The results are well-defined, and here is a pseudocode description of the edge cases. In this procedure, we'll use two internal variables named pending and detail.

1. Set pending to None.
2. Attempt to execute block $B_0$. If this block raises an exception $E_0$ with detail $d_0$, set pending to $E_0$ and set detail to $d_0$.
3. If pending is None, go to Step 8 (p. 84).
4. Find the first block except $E_i, v_i$: such that issubclass($E_0, E_i$).
   If there is no such match, go to Step 10 (p. 84).
5. Set $v_i$ to detail.
6. Attempt to execute block $B_i$.
   If this block raises some new exception $E_n$ with detail $d_n$, set pending to $E_n$ and set detail to $d_n$.
   However, if block $B_i$ executes without exception, set pending to None. In this case, the original exception is said to have been caught or handled.
7. Go to Step 10 (p. 84).
8. If there is no else: clause, go to Step 10 (p. 84).
9. Attempt to execute the else: block $B_e$.
   If this block raises some new exception $E_n$ with detail $d_n$, set pending to $E_n$ and set detail to $d_n$.
10. If there is no finally: clause, proceed to Step 12 (p. 85).
11. Attempt to execute the finally: block $E_f$. 
If this block raises some new exception $E_n$ with detail $d_n$, set pending to $E_n$ and set detail to $d_n$.

12. If pending is not None, re-raise the exception as in this statement:

```
raise pending, detail
```

If pending is None, fall through to the statement following the try: block.

### 16.9. The yield statement: Generate one result from a generator

A generator is any function or method that contains at least one yield statement. Generators are a special type of iterator; see Section 17.3, “Generators: Functions that can produce a sequence of values” (p. 89). Here is the general form:

```
yield expression
```

Unlike ordinary functions or methods that use the return statement to return a single value, a generator is a mechanism that produces a sequence of zero or more values. Each execution of a yield statement produces an additional value. To signal the caller that there are no more values, use this raise statement:

```
raise StopIteration
```

As an example, here is a function that generates the sequence $0, 1, 2, ..., n-1, n, n-1, n-2, ..., 2, 1, 0$.

```python
>>> def updown(n):
...     '''Generate the values 0, 1, 2, ..., n-1, n, n-1, n-2, ...0.
...     '''
...     for answer in range(0,n):
...         yield answer
...     for answer in range(n, -1, -1):
...         yield answer
...     raise StopIteration
...
>>> for x in updown(4):
...     print x,
... 0 1 2 3 4 3 2 1 0
```

### 17. def(): Defining your own functions

The `def` construct is used to define functions and methods. Here is the general form:

```
def n($p_0[=e_0][,p_1[=e_1]]...[,*p_v][,**p_d]):
    B
```

The name $n$ of the function is followed by a pair of parentheses containing descriptions of the arguments to the function. The block $B$ is called the body of the function, and is executed when the function is called.

A function may have no arguments at all. If there are arguments to be passed to the function when it is called, they must be declared in this order:

- A **positional argument** is a name that is not followed by an equal sign (=) and default value.
- A **keyword argument** is followed by an equal sign and an expression that gives its default value.
If a function has both positional arguments and keyword arguments, all positional arguments must precede all keyword arguments.

- If there is a `*p` parameter, when the function is called that name is bound to a (possibly empty) tuple of all positional arguments passed to the function that do not correspond to other positional or keyword arguments in the `def`.

- If there is a `**p` parameter, when the function is called that name is bound to a dictionary of all keyword arguments passed to the function that do not appear in the function's `def`.

When you call a function, the argument values you pass to it must obey these rules:

- There are two kinds of arguments: positional (also called non-default arguments) and keyword (also called default arguments). A positional argument is simply an expression, whose value is passed to the argument.

  A keyword argument has this form:

  ```
  name=expression
  ```

- All positional arguments in the function call (if any) must precede all keyword arguments (if any).

```
>>> def wrong(f=1, g):
...     print f, g
...
File "<stdin>", line 1
SyntaxError: non-default argument follows default argument
```

- You must supply at least as many positional arguments as the function expects.

```
>>> def wantThree(a, b, c):
...     print a,b,c
...
>>> wantThree('nudge', 'nudge', 'nudge')
nudge nudge nudge
>>> wantThree('nudge')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
TypeError: wantThree() takes exactly 3 arguments (1 given)
```

- If you supply more positional arguments than the function expects, the extra arguments are matched against keyword arguments in the order of their declaration in the `def`. Any additional keyword arguments are set to their default values.

```
>>> def f(a, b, c=1, d='elk'):
...     print a,b,c,d
...
>>> f(99, 111)
99 111 1 elk
>>> f(99, 111, 222, 333)
99 111 222 333
>>> f(8, 9, 10, 11, 12, 13)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
TypeError: f() takes at most 4 arguments (6 given)
```
• You may supply arguments for keyword parameters in any order by using the form \( k=v \), where \( k \) is the keyword used in the declaration of that parameter and \( v \) is your desired argument.

```python
>>> def blackKeys(fish='Eric', dawn='Stafford', attila='Abdul'):
...     print fish, dawn, attila
...     
>>> blackKeys()
Eric Stafford Abdul
>>> blackKeys(attila='Gamera', fish='Abdul')
Abdul Stafford Gamera
```

• If you declare a parameter of the form “*name”, the caller can provide any number of additional keyword arguments, and the name will be bound to a tuple containing those additional arguments.

```python
>>> def posish(i, j, k, *extras):
...     print i, j, k, extras
...     
>>> posish(38, 40, 42)
38 40 42 ()
>>> posish(44, 46, 48, 51, 57, 88)
44 46 48 (51, 57, 88)
```

• Similarly, you may declare a final parameter of the form “**name”. If the caller provides any keyword arguments whose names do not match declared keyword arguments, that name will be bound to a dictionary containing the additional keyword arguments as key-value pairs.

```python
>>> def extraKeys(a, b=1, *c, **d):
...     print a, b, c, d
...     
>>> extraKeys(1,2)
1 2 () {}
>>> extraKeys(3,4,6,12, hovercraft='eels', record='scratched')
3 4 (6, 12) {'record': 'scratched', 'hovercraft': 'eels'}
```

17.1. A function's local namespace

Any name that appears in a function's argument list, or any name that is set to a value anywhere in the function, is said to be local to the function. If a local name is the same as a name from outside the function (a so-called global name), references to that name inside the function will refer to the local name, and the global name will be unaffected. Here is an example:

```python
>>> x = 'lobster'
>>> y = 'Thermidor'
>>> def f(x):
...     y = 'crevettes'
...     print x, y
...     
>>> f('spam')
spam crevettes
>>> print x, y
lobster Thermidor
```
Keyword parameters have a special characteristic: their names are local to the function, but they are also used to match keyword arguments when the function is called.

### 17.2. Iterators: Values that can produce a sequence of values

Closely related to Python’s concept of sequences is the concept of an *iterator*:

For a given sequence $S$, an iterator $I$ is essentially a set of instructions for producing the elements of $S$ as a sequence of zero or more values.

To produce an iterator over some sequence $S$, use this function:

```
iter(S)
```

- The result of this function is an “iterator object” that can be used in a `for` statement.

```python
>>> continents = ('AF', 'AS', 'EU', 'AU', 'AN', 'SA', 'NA')
>>> worldWalker = iter(continents)
>>> type(worldWalker)
<type 'tupleiterator'>
>>> for landMass in worldWalker:
...    print "Visit %s." % landMass,
...Visit AF. Visit AS. Visit EU. Visit AU. Visit AN. Visit SA. Visit NA.
```

- An iterator must have a `.next()` method that you can call to get the next element in the sequence. This method takes no arguments. It returns the next element in the sequence, if any. When there are no more elements, it raises a `StopIteration` exception.

```python
>>> trafficSignal = [ 'green', 'yellow', 'red' ]
>>> signalCycle = iter(trafficSignal)
>>> type(signalCycle)
<type 'listiterator'>
>>> signalCycle.next()
'green'
>>> signalCycle.next()
'yellow'
>>> signalCycle.next()
'red'
>>> signalCycle.next()
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
StopIteration
```

Once an iterator is exhausted, it will continue to raise `StopIteration` indefinitely.

- You can also use an iterator as the right-hand operand of the “in” operator.

```python
>>> 'red' in signalCycle
True
```
17.3. Generators: Functions that can produce a sequence of values

Unlike conventional functions that return only a single result, a **generator** is a function that produces a sequence of zero or more results.

Generators are a special case of iterators (see Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88)), so they can be used as the controlling sequence in `for` statements and the other places where iterators are allowed.

In a conventional function, the body of the function is executed until it either executes a `return` statement, or until it runs out of body statements (which is the equivalent of a “`return None`” statement).

By contrast, when a generator function is called, its body is executed until it either has another value to produce, or until there are no more values.

- When a function wishes to return the next generated value, it executes a statement of this form:

  ```python
  yield e
  ```

  where the `e` is any Python expression.

  The difference between `yield` and `return` is that when a `return` is executed, the function is considered finished with its execution, and all its current state disappears.

  By contrast, when a function executes a `yield` statement, execution of the function is expected to resume just after that statement, at the point when the caller of the function needs the next generated value.

- A generator signals that there are no more values by executing this statement:

  ```python
  raise StopIteration
  ```

For an example of a generator, see Section 16.9, “The `yield` statement: Generate one result from a generator” (p. 85).

If you are writing a container class (that is, a class whose instances are containers for a set of values), and you want to define an iterator (see Section 19.3.14, “`__iter__()`: Create an iterator” (p. 107)), that method can be a generator. Here is a small example. The constructor for class **Bunch** takes a sequence of values and stores them in instance attribute `__stuffList`. The iterator method `__iter__()` generates the elements of the sequence in order, except it wraps each of them in parentheses:

```python
>>> class Bunch(object):
...     def __init__(self, stuffList):
...         self.__stuffList = stuffList
...     def __iter__(self):
...         for thing in self.__stuffList:
...             yield "(%s)" % thing
...         raise StopIteration
... >>> mess = Bunch(('lobster Thermidor', 'crevettes', 'Mornay'))
... >>> for item in mess:
...     print item,
... (lobster Thermidor) (crevettes) (Mornay)
... >>> messWalker = iter(mess)
... >>> for thing in messWalker: print thing,
... (lobster Thermidor) (crevettes) (Mornay)
```
18. Exceptions: Error signaling and handling

Python’s exception system provides a way to signal error conditions and other disruptions in normal processing, and also a way for programs to recover from these conditions.

• Section 18.1, “Definitions of exception terms” (p. 90).
• Section 18.2, “Life cycle of an exception” (p. 91).
• Section 18.3, “Built-in exceptions” (p. 91).

18.1. Definitions of exception terms

Some definitions:

• To raise an exception means to signal that the program cannot proceed normally due to an error or other condition. (In other programming languages, such as Java, the equivalent term is to throw an exception.)

Two values accompany the raising of an exception: the type and the value. For example, if a program attempts to open an existing disk file but there is no such file, the type is IOError, and the value is an instance of the IOError class that contains additional information about this error.

For more information about raising exceptions, see Section 16.6, “The raise statement: Cause an exception” (p. 81).

• A program may choose to handle an exception. That is, a program may say that if a certain exception or category of exceptions occurs in a specific block of code, Python must execute another code block called a handler.

• A traceback is a message from Python showing where an exception occurred.

If you type a statement in conversational mode that causes an exception, you will see a short traceback like this:

```python
>>> x = 59 / 0
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ZeroDivisionError: integer division or modulo by zero
```

The above example showed that the offending statement was read from the standard input stream (<stdin>).

When looking at a traceback, always look at the last line first. It tells you the general type of exception (in the example, a ZeroDivisionError), followed by additional details (“integer division or modulo by zero”).

If an exception occurs inside one or more function calls, the traceback will give a complete list of the functions involved, from outermost to innermost. Again, the last line shows the exception type and details.

```python
>>> def f():
...     g()
... >>> def g():
...     h()
... >>> def h():
...     return 1/0
... >>> f()
Traceback (most recent call last):
```

18.2. Life cycle of an exception

If you anticipate that executing a particular statement may cause an exception and you don’t want your program to terminate and display a traceback, you can use a `try` construct to specify handlers to be executed if an exception occurs. For details, see Section 16.8, “The try statement: Anticipate exceptions” (p. 82).

If an exception occurs inside a function and it is not handled at that level by a `try` construct, Python will work back through the pending function calls until it either finds a handler for that exception or runs out of pending function calls.

If there are multiple handlers for the exception in calling functions, the innermost will be used. If there are no handlers for the exception in calling functions, you will get a stack traceback and the program will terminate.

```python
>>> def f():
...     try:
...         g()
...     except ValueError, detail:
...         print "Caught a ValueError: ", detail.message
... >>> def g(): h()
... >>> def h():
...     raise ValueError('This is a test.')
... >>> f()
Caught a ValueError: This is a test.
```

In the example above, function `f()` calls function `g()`, which in turn calls function `h()`. Function `h()` raises a `ValueError` exception, but there is no `try:` block around it. Python looks to see if there is a `ValueError` handler in `g()`, but there is not. Finally a handler for `ValueError` is found inside function `f()`, so control resumes inside that handler. Note that no stack traceback is displayed, because the `ValueError` exception was handled successfully.

18.3. Built-in exceptions

Python defines a complete hierarchy of built-in exception classes. When you write a handler, you can specify any class in this hierarchy, and that handler will apply to that class and any derived classes. This allows you to write generic handlers that catch whole groups of exception types.

In describing this hierarchy, we will use indentation to show the subclass/parent class relationships. Generally a `<raise>` statement will name one of the “leaf” classes, that is, a class that does not have any subclasses. Ancestor classes that are not usually raised are marked with an asterisk (*) in the section below.

---

New Mexico Tech Computer Center  Python 2.5 quick reference  91
- **BaseException**: This is the ancestor of all exception classes. The constructor for this class takes one argument, a string containing the error message. That argument is available as the `.message` attribute.

- **SystemExit**: Raise this exception to terminate execution of your program; it is a special case in that it does not produce a stack traceback. In module `sys`, the `.exit()` method raises this exception. It is possible to write a handler for this exception. To defeat such a handler and force immediate termination, import module `os` and use method `os._exit()`.

- **KeyboardInterrupt**: Raised when the user signals an interruption with the keyboard (`del` under Windows, or `Control-C` in Linux or Mac environments). This class inherits from `BaseException` rather than from `Exception` so that catch-all handlers for class `Exception` will not prevent program termination.

- **Exception**: This is the preferred base class for all built-in and user-defined exceptions. If you want to write a handler for all these exception types, use `Exception` as the handler’s type.

```python
>>> try:
...     x = 1 / 0
... except Exception, detail:
...     print "Fail:\n", detail.message
... Fail: integer division or modulo by zero
>>> try:
...     x = noSuchVariable
... except Exception, detail:
...     print "Fail:\n", detail.message
... Fail: name 'noSuchVariable' is not defined
```

**Warning**

A catch-all handler like this can mask any number of errors. Do not use such a handler unless your program must absolutely stay running.

- **StopIteration**: This is the exception that a generator must raise in order to signal that no more generated values are available. See Section 17.3, “Generators: Functions that can produce a sequence of values” (p. 89).

- **StandardError**: This is the base class for all built-in exceptions that are considered errors.

  - **ArithmeticError**: This is the base class for errors involving arithmetic computations.
    - **FloatingPointError**: This is raised for arithmetic errors involving the `float` type.
    - **OverflowError**: This is raised when the result of an operation cannot be represented.
    - **ZeroDivisionError**: An attempt to divide by zero.

  - **AssertionError**: An `assert` statement has failed. See Section 15.2, “The `assert` statement: Verify preconditions” (p. 72).

  - **AttributeError**: Failure to access an attribute.

  - **EnvironmentError**: Errors caused by functions outside of Python, such as the operating system or peripheral devices.
• IOError: Errors related to file input or output.
• OSError: Errors signaled from the operating system.
• ImportError: Failure to import a module or to import items from a module.
• LookupError*: Superclass for errors caused by attempts to retrieve values from inside a container class.
• IndexError: Attempt to retrieve a sequence member \( S[I] \), where \( I \) is not a valid index in sequence \( S \).
• KeyError: Attempt to retrieve a dictionary member \( D[K] \), where \( K \) is not a valid key in \( D \).
• MemoryError: No more processor memory is available.
• NameError: Attempt to retrieve a name that is not defined.
• UnboundLocalError: Attempt to retrieve the value of a local name when no value has yet been assigned to it.
• RuntimeError: An error that doesn’t fit the other categories.
• NotImplementedError: This is the preferred way for the virtual methods of a base class to signal that they have not been replaced by a concrete method in a derived class.
• SyntaxError: Attempt to execute invalid Python source code.
• TypeError: Attempt to perform an operation on a value that does not support that operation, such as trying to use exponentiation (***) on a string.
• ValueError: Caused by an operation that is performed on values of the correct type, but the actual values are not valid. Example: taking a negative number to a fractional power.

19. Classes: Defining your own types

This section assumes you already understand the basics of object-oriented programming in Python, and that you know the meaning of concepts such as class, instance, method, and attribute. For a general tutorial on these concepts, see the introduction to object-oriented Python programming\(^{12}\) in the Tech Computer Center’s Python tutorial\(^{13}\).

Here is the general form of the class declaration for some class \( C \) with one or more parent classes \( P_1, P_2, \ldots \):

```python
class C(P_1, P_2, ...):
    attribute definitions
    ...
```

To declare a class that does not inherit from any parent classes:

```python
class C:
    attribute definitions
    ...
```

The attribute definitions may include any number of def blocks that declare methods of the class, and any number of class variable declarations.

\(^{13}\) http://www.nmt.edu/tcc/help/pubs/lang/pytut/
Functionally, a class is really just a namespace. This namespace is just a place to store the pieces of the class mechanisms: its methods and class variables.

- When Python reads a “class” declaration, it creates a new, empty namespace.
- When Python reads a “def” within a class, the name of that method is added to the class’s namespace.
- If you define a class variable (that is, if you assign a value to a name inside a class but outside of any methods of the class), the class variable’s name and value are added to the namespace.

A brief conversational session may serve to illustrate these concepts. We’ll make use of the built-in function `dir()` to show the contents of the class’s namespace; see Section 14.5, “`dir()`: Display a namespace’s names” (p. 60).

```python
>>> class Taunter:
...     tauntCount = 0
...     def taunt(self):
...         print "Go away, or I shall taunt you a second time!"
...
>>> dir(Taunter)
['__doc__', '__module__', 'taunt', 'tauntCount']
>>> type(Taunter.__doc__)
<type 'NoneType'>
>>> Taunter.__module__
'__main__'
>>> Taunter.tauntCount
0
>>> Taunter.taunt
<unbound method Taunter.taunt>
```

1. When Python reads this line, it adds the name `Taunter` to the current local namespace, bound to a new, empty namespace of type `class`.
2. Because this assignment takes place inside class `Taunter` but not inside a `def`, name `tauntCount` becomes a class variable, bound to an `int` value of zero.
3. The next two lines define a method named `taunt()` within the class.
4. After we’ve finished entering the class definition, we use `dir(Taunter)` to see what names are in the class’s namespace. Variables `__doc__` and `__module__` are added automatically. Because there was no documentation string in the class, `__doc__` is bound to `None`. The `__module__` variable has the value `['__main__']` because the class was entered in conversational mode.
5. To retrieve the value of a class variable `V` in class `C`, use the syntax “`C.V`”.
6. Name `taunt` in the class namespace is bound to an object of type “unbound method.” An unbound method is a method (function) that is inside a class, but it is not associated with an instance of the class.

An instance of a class is also a namespace. When the instance is created, all the names from the class’s namespace are copied into the instance namespace. From that point on, any changes made to the instance’s namespace do not affect the class namespace:

```python
>>> frenchy = Taunter()
>>> dir(frenchy)
['__doc__', '__module__', 'taunt', 'tauntCount']
>>> frenchy.where = 'crenelations'
>>> dir(frenchy)
['__doc__', '__module__', 'where', 'taunt', 'tauntCount']
>>> frenchy.where

94
This class does not have a constructor (_init_) method, so when an instance is created, the instance is a namespace with the same names as the class, and the same values.

This line adds a new name where to the instance’s namespace. It is bound to the string value 'crenelations'.

Note that the instance namespace now contains the name where, but the class’s namespace is unchanged.

To retrieve an attribute A of an instance I, use the syntax “I.A”. Initially, the instance variable has the same value as the class variable of the same name.

Here, we add one to the instance variable tauntCount. The instance variable has the new value, but the class variable tauntCount is unchanged.

Within the instance namespace, name taunt is now a bound method: it is associated with the instance frenchy.

The next two lines show two equivalent methods of calling the taunt method.

Most method calls are bound method calls. To call a bound method B of an instance I, use the syntax “I.B(...)”.

When a method B is bound to an instance I, the instance namespace I becomes the “self” argument passed in to the method.

This line has the same effect as the previous line, but it is an unbound method call.

The expression “Taunter.taunt” retrieves the unbound method from the class definition. When you call an unbound method, you must supply the “self” argument explicitly as the first argument.

Unbound method calls are not terribly common, but you will need to know about them when you write the constructor for a derived class: you must call the parent class constructor as an unbound call. Generally, if class D has parent class C, the derived class might look something like this:

```python
class D(C):
    def __init__(self, ...):
        C.__init__(self, ...)
        ...
```

Namespaces are very much like dictionaries. Where a dictionary has unique keys, a namespace has unique names. As a matter of fact, classes and instances have a special built-in attribute called “__dict__” which, for most purposes, is the namespace as a dictionary. Continuing the examples above:
The class's dictionary has the four names we expect: the built-ins __module__ and __doc__, the class variable tauntCount, and the method taunt.

But notice that the __dict__ attribute of the newly created instance newFrenchy does not have the four names copied from the class. In fact, it is empty. And the __dict__ of instance frenchy contains only the names that have changed since its instantiation.

What actually happens when you refer to an attribute is that Python looks first in the instance's __dict__; if the name is not found there, it looks in the __dict__ of the class. For derived classes, Python will also search the __dict__ attributes of all the ancestor classes.

So, in our example, a reference to frenchy.tauntCount would find the value of 1 in the instance. A reference to newFrenchy.tauntCount would fail to find that name in newFrenchy.__dict__, but would succeed in finding the class variable value 0 in Taunter.__dict__['tauntCount'].

Let's now look at the life cycles of classes in more detail. Due to improvements made in the language since it was first introduced, Python has two kinds of classes, old-style and new-style. We encourage you to use new-style classes; old-style classes will no longer be supported in the next major release, Python 3000.

• All new-style classes must declare at least one parent class that is either the top-level class object or some other class that derives ultimately from object. Such a class is said to be derived from, or inherits from, the object class.

To declare a new-style class C that inherits from object:

```python
class C(object):
    ...
class methods and variables...
```

• An old-style class is one that doesn’t declare a parent class at all, or a class that inherits from an existing old-style class. The life cycle of an old-style class is described in Section 19.1, “Old-style classes” (p. 96).

In most respects, the two classes perform identically.

• We'll start by explaining old-style classes in Section 19.1, “Old-style classes” (p. 96).
• To benefit from the many functional improvements of new-style classes, and especially if you expect to migrate your code to the major changes of Python 3.0, see Section 19.2, “Life cycle of a new-style class” (p. 99).

### 19.1. Old-style classes

Old-style classes are those declared without a parent class, or classes that inherit from an existing old-style class.

Here is an outline of the birth, life, and death of an old-style class and its instances.

• Section 19.1.1, “Defining an old-style class” (p. 97).
• Section 19.1.2, “Instantiation of an old-style class: The constructor, .__init__()” (p. 97).
19.1.1. Defining an old-style class

To define an old-style class \( C \) with no parent class, use this general form:

```python
class C:
    ...class methods and variables...
```

To create a class that inherits from one or more parent classes \( P_1, P_2, \ldots \):

```python
class C(P_1, P_2, ...):
    ...class methods and variables...
```

As Python reads the definition of your class, it first creates a new, empty namespace called the class namespace. You can access the class namespace directly as an attribute named `__dict__`, a dictionary whose keys are the names in that namespace.

As Python reads each method or class variable in your class declaration, it adds the name of that method or variable to the class namespace.

19.1.2. Instantiation of an old-style class: The constructor, \().__init__()\)

The creation of a new instance of a class happens when a running Python program encounters a call to that class, that is, the class name with a pair of parentheses after it, with zero or more arguments inside the parentheses.

Here is the general form:

```python
C(a_1, a_2, ...)
```

The instance creation (also called instantiation) is handled by the `__init__()` or constructor method.

1. First Python creates the instance with an empty `__dict__` attribute that will contain the instance's values.
2. Python then calls the constructor. The argument list for this call always has the special first argument `self` (the instance), followed by whatever arguments were used in the initial call. The constructor call is equivalent to this:

```python
C.__init__(self, a_1, a_2, ...)
```

3. The constructor method then executes. Typically the constructor will set up new instance attributes by assignments of this form:

```python
self.name = expression
```

When the constructor finishes executing, the instance is returned to the constructor's caller.

19.1.3. Attribute references in old-style classes

The names inside the instance are called attributes. (The methods are technically attributes—attributes that are of type function.) There are three operations on attributes: get, set, and delete.
To get an attribute means to retrieve its value. Python searches the instance’s .__dict__ for the attribute’s name; if that name is found, its value is returned. If the instance’s .__dict__ does not have a binding for the given name, Python searches the class’s .__dict__. If no value is found there, Python searches the namespaces of the ancestor classes (if any).

```python
g>>> class C:
...     def __init__(self, x):
...         self.thingy = x
...>>> c=C(42)
>>> c.thingy
42
>>> c.__dict__['thingy']
42
```

When you call a method \( M \) of an instance \( I \) in an expression of the form “\( I.M(...) \)”, this is considered just another attribute “get” operation: the get operation \( I.M \) retrieves the method, and then that method is called using the arguments inside the “(...)”.

To set an attribute means to give it a value. If there is an existing attribute with the same name, its old value is discarded, and the attribute name is bound to the new value. The new value is stored in the instance's .__dict__.

```python
g>>> c.thingy
42
>>> c.thingy = 58
>>> c.thingy
58
>>> c.__dict__['thingy']
58
```

You can delete an attribute from an instance using a del statement (see Section 15.3, “The del statement: Delete a name or part of a value” (p. 73)).

```python
g>>> c.thingy
58
>>> del c.thingy
>>> c.thingy
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  AttributeError: C instance has no attribute 'thingy'
```

In addition to ordinary attributes and methods, your class can accept references to names that do not exist in the instance or class namespace. You can define special methods that will be called when some statement tries to get, set, or delete an attribute that isn't found in the instance's .__dict__. See Section 19.3.11, “__getattr__()” (p. 106), Section 19.3.17, “__setattr__()” (p. 108), and Section 19.3.9, “__delattr__()” (p. 106).

If all else fails—if an attribute is not found in the instance's namespace and the class does not provide a special method that handles the attribute reference—Python will raise an AttributeError exception.

19.1.4. Method calls in an old-style class

There are two different ways to call a method \( M \) of some class \( C \):
Most calls are *bound method* calls of this form, where *I* is an instance of some class *C*:

```
I.methodName(a_1, a_2, ...)
```

The instance *I* replaces `self` as the first argument when setting up the arguments to be passed to the method.

The following form, called an *unbound method call*, is exactly equivalent to the above:

```
C.methodName(i, a_1, a_2, ...)
```

Here is a demonstration of the equivalence of bound and unbound method calls.

```python
class C:
    ... def __init__(self, x): self.x = x
    ... def show(self, y): print "*** (%s,%s) ***" % (self.x, y)
    ...
c=C(42)
c.show(58)
>>> C.show(c,58)
*** (42,58) ***
```

### 19.1.5. Instance deletion: the destructor, `__del__()`

When there are no values that refer to an instance any more, the storage occupied by the instance is recycled. However, if there are certain cleanup operations that must be done (such as closing external files), these operations can be placed into a *destructor* method that will be called before recycling the instance. Here is the general form of a destructor:

```
def __del__(self):
    ...
cleanup statements...
```

### 19.2. Life cycle of a new-style class

Most of the features of new-style classes are the same as for old-style classes. This section will discuss only the differences. We won’t cover a few of the more obscure advanced features here; for information on such topics as descriptors and metaclasses, see the “Data model” section of the *Python Reference Manual*.

The declaration of a new-style class looks the same as for an old-style class, with one constraint: the class must inherit from the universal base class named `object`, or from one or more other new-style classes.

- Section 19.2.1, “`__new__()`: New instance creation” (p. 100).
- Section 19.2.2, “Attribute access control in new-style classes” (p. 100).
- Section 19.2.4, “Conserving memory with `__slots__`” (p. 101).

---

14 [http://docs.python.org/ref/datamodel.html](http://docs.python.org/ref/datamodel.html)
19.2.1. __new__(): New instance creation

New-style classes have a new special method name, __new__(), that is called on instantiation before the constructor. It handles the creation of a new instance.

- The __new__() method is called when an instance is created.
- Method __new__() is always a static method (see Section 19.4, “Static methods” (p. 108)), even if you do not specifically make it a static method.
- A constructor call for some class C has this general form:

\[
C(*p, **k)
\]

That is, it can have any number of positional arguments and any number of keyword arguments.

The equivalent call to the __new__() method will look like this:

```python
def __new__(cls, *p, **k):
    ...
```

The first argument cls must be the class being created.

- The __new__() method must call the parent class's __new__() method to create the instance. For example, if your class inherits directly from object, you must call:

```python
object.__new__(cls)
```

The value returned by that call is the new instance.

- In most cases, the __new__() method will return a new instance of cls, and that class's __init__() will then be called with that instance as its self argument, and the positional and keyword arguments p and k will be passed to that constructor as well.

```python
>>> class Test(object):
...     def __new__(cls, *p, **k):
...         inst = object.__new__(cls)
...         return inst
...     def __init__(self, *p, **k):
...         print "p=%s k=%s" % (p, k)
...     ...

>>> t=Test('egg', 'kale', sauce='Bearnaise')
p=('egg', 'kale') k={'sauce': 'Bearnaise'}
```

However, if the __new__() method does not return an instance of class cls, the constructor method __init__() will not be called. This allows the class more control over how new instances are created and initialized. You can return an instance of an entirely different class if you like.

19.2.2. Attribute access control in new-style classes

New-style classes give you more ways to control what happens when an instance's attribute is accessed.

Here is the general procedure for access to attribute a of instance i, where C is the class of i.

1. If the instance has a __getattr__() special method (see Section 19.3.12, “__getattr__() Special Method” (p. 107)), execute that method, which must either return the attribute value or raise AttributeError.
2. If the instance has a \_\_slots\_ attribute (see Section 19.2.4, “Conserving memory with \_\_slots\_” (p. 101)), return the value of the slot with name \(a\). If \(a\) does not match any of the slot names, or if the named slot has never been set to a value, raise AttributeError.

3. If \(a\) is a key in \_\_dict\_, return the corresponding value.

4. Search for attribute \(a\) in class \(C\). If that fails, search all the parent classes of \(C\) all the way back to object.

5. If all searches in the preceding step failed and the instance has a \_\_getattr\_\_ special method, call that method. See Section 19.3.11, “\_\_getattr\_\_” (p. 106); please note the differences from Section 19.3.12, “\_\_getattribute\_\_” (p. 107).

6. If all the above steps fail to produce a value, raise AttributeError.

19.2.3. Properties in new-style classes: Fine-grained attribute access control

The outline of attribute access in Section 19.2.2, “Attribute access control in new-style classes” (p. 100) is slightly oversimplified in one respect. Any of the attribute search steps in this procedure may produce a property rather than the actual attribute value.

A property is a special object that is produced by the property() function. For a discussion of the three types of attribute access (get, set, and delete), the protocols for the accessor functions, and examples, see Section 14.15, “property()” (p. 65).

19.2.4. Conserving memory with \_\_slots\_

Normally, you can add new attributes to an instance’s namespace with any name you want. The instance’s \_\_dict\_ attribute is effectively a dictionary, and you can add any number of names to it.

However, in a new-style class, you may specify a given, limited set of attribute names that are allowed in instances of the class. There are two reasons why you might want to do this:

- If your program is going to create large numbers of instances of a class, to the point where you may run out of memory, you can save some storage within each instance by sacrificing the ability to add arbitrary attribute names.
- If you limit the set of permissible attribute names, Python will detect any reference to a name not in the permissible set, and raise an AttributeError exception. This may help you catch certain programming errors.

To limit the set of attribute names in a new-style class, assign to a class variable named \_\_slots\_ a tuple containing the allowable names, like this:

\[
\_\_slots\_ = (n_1, n_2, \ldots)
\]

Here’s a small example. Suppose you want instances of class Point to contain nothing more than two attributes named .\(x\) and .\(y\):

```python
>>> class Point(object):
...     __slots__ = ('x', 'y')
...     def __init__(self, abscissa, ordinate):
...         self.x, self.y = abscissa, ordinate
...     ...
>>> x2=Point(3, 7)
>>> x2.x
```

New Mexico Tech Computer Center Python 2.5 quick reference
When you declare a `__slots__` attribute in a new-style class, instances will *not* have a `__dict__` attribute.

### 19.3. Special method names

Within a class, a number of reserved method names have special meaning.

The most important special method name is the class constructor, `__init__()`.

- For a general introduction to class constructors, see Section 19, “Classes: Defining your own types” (p. 93).
- For old-style class constructors, see Section 19.1.2, “Instantiation of an old-style class: The constructor, `__init__()`” (p. 97).
- For new-style class constructors, see Section 19.2, “Life cycle of a new-style class” (p. 99).

Many special methods fall into broad categories:

- Section 19.3.1, “Rich comparison methods” (p. 103): for methods that implement comparisons such as “<” and “==”.
- Section 19.3.2, “Special methods for binary operators” (p. 103): for operators that operate on two operands, such as “%”.
- Section 19.3.3, “Unary operator special methods” (p. 104): for operators that operate on a single operand, such as negation, “-”.
- Section 19.3.4, “Special methods to emulate built-in functions” (p. 104): for classes that handle calls to built-in functions such as “str()”.

The remaining special methods will be discussed in alphabetical order.

- Section 19.3.5, “__call__()”: What to do when someone calls an instance” (p. 105).
- Section 19.3.6, “__cmp__()”: Generalized comparison” (p. 105).
- Section 19.3.8, “__del__()”: Destructor” (p. 106).
- Section 19.3.9, “__delattr__()”: Delete an attribute” (p. 106).
- Section 19.3.10, “__delitem__()”: Delete one item of a sequence” (p. 106).
- Section 19.3.11, “__getattr__()”: Handle a reference to an unknown attribute” (p. 106).
- Section 19.3.12, “__getattribute__()”: Intercept all attribute references” (p. 107).
- Section 19.3.13, “__getitem__()”: Get one item from a sequence or mapping” (p. 107).
- Section 19.3.14, “__iter__()”: Create an iterator” (p. 107).
- Section 19.3.15, “__nonzero__()”: True/false evaluation” (p. 107).
- Section 19.3.16, “__repr__()”: String representation” (p. 107).
- Section 19.3.17, “__setattr__()”: Intercept all attribute changes” (p. 108).
- Section 19.3.18, “__setitem__()”: Assign a value to one item of a sequence” (p. 108).
19.3.1. Rich comparison methods

These special methods allow your class to specify what happens when comparison operators such as “<=” are used, and the left-hand operator is an instance of your class. (In most cases the right-hand operator is also an instance of the same class, but this is not actually required.)

In each case, the calling sequence for the method must look like this:

```python
def __method__(self, other):
    ...
```

The `self` argument is the left-hand operand and the `other` argument is the operand on the right hand of the operator.

Each method must return a numeric value:

- A negative number indicates that `self` precedes `other`.
- Zero indicates that `self` and `other` are considered equal.
- A positive number indicates that `other` precedes `self`.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Method name</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>==</code></td>
<td><code>__eq__</code></td>
</tr>
<tr>
<td><code>&gt;=</code></td>
<td><code>__ge__</code></td>
</tr>
<tr>
<td><code>&gt;</code></td>
<td><code>__gt__</code></td>
</tr>
<tr>
<td><code>&lt;=</code></td>
<td><code>__le__</code></td>
</tr>
<tr>
<td><code>&lt;</code></td>
<td><code>__lt__</code></td>
</tr>
<tr>
<td><code>!=</code></td>
<td><code>__ne__</code></td>
</tr>
</tbody>
</table>

19.3.2. Special methods for binary operators

Your class can define special methods with these names to tell Python how to handle binary operators such as “*” or “%”. In each case, the calling sequence will look something like this:

```python
def __method__(self, other):
    ...
```

The `self` argument is the left-hand operand, and the `other` argument is the right-hand operand. Your method will return the result of the operation.

For each operator, you may supply up to three methods:

- The method in the first column performs the normal operation.
- The method in the second column is used when the left-hand operand does not support the given operation and the operands have different types. In these methods, `self` is the right-hand operand and `other` is the left-hand operand.
- The third column implements the “augmented assignment” operators such as “+=”. For example, for method `__iadd__(self, other)`, the method must perform the equivalent of “`self += other`”.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Normal</th>
<th>Reversed</th>
<th>Augmented</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>+</code></td>
<td><code>__add__</code></td>
<td><code>__radd__</code></td>
<td><code>__iadd__</code></td>
</tr>
<tr>
<td><code>&amp;</code></td>
<td><code>__and__</code></td>
<td><code>__rand__</code></td>
<td><code>__iand__</code></td>
</tr>
<tr>
<td><code>/</code></td>
<td><code>__div__</code></td>
<td><code>__rdiv__</code></td>
<td><code>__idiv__</code></td>
</tr>
</tbody>
</table>
19.3.3. Unary operator special methods

You can define these special methods in your class to specify what happens when a unary operator such as "-" (negate) is applied to an instance of the class.

In each case, the definition will look like this:

```python
def __method__(self):
    ...
```

The method returns the result of the operation.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>~</td>
<td><strong>invert</strong></td>
</tr>
<tr>
<td>-</td>
<td><strong>neg</strong></td>
</tr>
<tr>
<td>+</td>
<td><strong>pos</strong></td>
</tr>
</tbody>
</table>

19.3.4. Special methods to emulate built-in functions

You can define special methods that will handle calls to some of Python's built-in functions. The number of arguments will be the same as for the built-in functions, except that `self` is always the first argument.

For example, a special method to handle calls to function `divmod(x, y)` will look like this:

```python
def __divmod__(self, other):
    ...
```

In this method, the value of the first argument will be passed to `self` and the second argument to `other`.

<table>
<thead>
<tr>
<th>Function</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>abs</td>
<td><strong>abs</strong></td>
</tr>
<tr>
<td>complex</td>
<td><strong>complex</strong></td>
</tr>
<tr>
<td>divmod</td>
<td><strong>divmod</strong></td>
</tr>
<tr>
<td>hex</td>
<td><strong>hex</strong></td>
</tr>
<tr>
<td>int</td>
<td><strong>int</strong></td>
</tr>
</tbody>
</table>
### 19.3.5. `__call__()`: What to do when someone calls an instance

If a class has a `__call__()` method, its instances can be called as if they were functions. Any arguments passed in that function call become arguments to the `__call__()` method. Example:

```python
class CallMe(object):
    def __call__(self, *p, **k):
        print "CallMe instance called with:"
        print "Positional arguments", p
        print "Keyword arguments", k

c=CallMe()
c(1, 'rabbit', fetchez='la vache', hamster='elderberries')
```

This will result in:

```
CallMe instance called with:
Positional arguments (1, 'rabbit')
Keyword arguments {'fetchez': 'la vache', 'hamster': 'elderberries'}
```

### 19.3.6. `__cmp__()`: Generalized comparison

The purpose of this special method is to implement comparison operations between instances. It will be called in these situations:

- If the built-in `cmp()` function is called to compare an instance of a class to some other value, and the class has a `__cmp__()` method, that method is called to perform the comparison.

- When an instance appears on the left-hand side of a comparison (relational) operator, and that instance's class has a the corresponding rich-comparison method (such as `__eq__()` for the `==` operator; see Section 19.3.1, “Rich comparison methods” (p. 103)), the rich-comparison method will be called to perform the comparison. and, if so, that method is called.

The comparison operators are "<", "<=", "==", "!=", ">", and ">=".

However, if the class does not have the correct rich-comparison method, but it does have a `__cmp__()` method, that method will be called to evaluate the comparison.

The calling sequence is:

```python
def __cmp__(self, other):
...```

The convention for return values is the same one described in Section 13.6, “`cmp()`: Compare two values” (p. 46): negative if `self` precedes `other`, positive if `other` precedes `self`, zero if they are considered equal.
19.3.7. __contains__(): The “in” and “not in” operators

This special method defines how instances of a class behave when they appear as the right-hand operand of Python’s “in” and “not in” operators.

Here is the calling sequence:

```python
def __contains__(self, x):
    ...
```

The method returns True if x is considered to be in self, False otherwise.

19.3.8. __del__(): Destructor

If a class has a .__del__() method, that method is called when an instance is deleted. For details, see Section 19.1.5, “Instance deletion: the destructor, .__del__()” (p. 99).

19.3.9. __delattr__(): Delete an attribute

If a class has a .__delattr__() method, that method is called when an attribute is deleted. Here is the calling sequence:

```python
def __delattr__(self, name):
    ...
```

The name argument is the name of the attribute to be deleted.

19.3.10. __delitem__(): Delete one item of a sequence

This method defines the behavior of a del statement of this form:

```python
del s[i]
```

Such a statement can be used either on objects that act like sequences, where i specifies the position of the element to be deleted, or mapping objects (that is, dictionary-like objects), where i is the key of the key-value pair to be deleted.

The calling sequence is:

```python
def __delitem__(self, i):
    ...
```

19.3.11. __getattr__(): Handle a reference to an unknown attribute

This method, if present, handles statements that get an attribute value of an instance, but there is no such entry in the instance’s namespace. The calling sequence is:

```python
def __getattr__(self, name):
    ...
```

The argument is the name of the desired attribute. The method must either return the attribute’s value or raise an AttributeError exception.

Compare Section 19.3.12, “__getattribute__(): Intercept all attribute references” (p. 107), which is called even if the instance namespace does have an attribute with the desired name.
19.3.12. __getattr__(): Intercept all attribute references

This method is called whenever an attribute is referenced, whether the instance or class namespace has an attribute with the given name or not. It works only with new-style classes.

For an overview and examples, see Section 19.2.2, “Attribute access control in new-style classes” (p. 100).

19.3.13. __getitem__(): Get one item from a sequence or mapping

If a class defines it, this special method is called whenever a value is retrieved from a sequence or mapping (dictionary-like object) using the syntax “v[i]”, where v is the sequence or mapping and i is a position in a sequence, or a key in a mapping.

Here is the calling sequence:

```python
def __getitem__(self, i):
...
```

The method either returns the corresponding item or raises an appropriate exception: IndexError for sequences or KeyError for mappings.

19.3.14. __iter__(): Create an iterator

If a class defines an .__iter__() method, that method is called:

• Whenever the built-in iter() function is applied to an instance of the class.

• In any situation where the instance is iterated over, such as when it appears as the controlling sequence of a for statement.

The calling sequence is:

```python
def __iter__(self):
...
```

The return value must be an iterator. An iterator is any object that has a .next() method that returns the next value in sequence, or raises StopIteration when the sequence is exhausted. For more information, see Section 17.2, “Iterators: Values that can produce a sequence of values” (p. 88).

19.3.15. __nonzero__(): True/false evaluation

If a class defines it, this special method is called whenever an instance is converted to a Boolean value, either implicitly (for example, when it is the test in an “if” statement) or explicitly via the built-in bool() function. Here is the calling sequence:

```python
def __nonzero__(self):
...
```

Return a Boolean value, either True or False.

19.3.16. __repr__(): String representation

If a class defines it, this special method is called to find the “representation” of an object. There are two ways to get a representation:

• Via the function described in Section 14.17, “repr() : Representation” (p. 66).
• Via the “back-quote operator”, enclosing a Python expression in open single quotes.

The calling sequence is:

```python
def __repr__(self):
    ...
```

The method returns the representation of `self` as a string.

### 19.3.17. `__setattr__()`: Intercept all attribute changes

If a class defines it, this method is called whenever a new value is stored into an attribute. Calling sequence:

```python
def __setattr__(self, name, value):
    ...
```

The method sets the attribute given by the `name` argument to the `value` argument, or raises `AttributeError` if that operation is not permitted.

### 19.3.18. `__setitem__()`: Assign a value to one item of a sequence

If a class defines it, this method is called whenever a new value is stored into a sequence or mapping (dictionary-like object), such as in statements of this form:

```python
V[i] = expr
```

Here is the calling sequence:

```python
def __setitem__(self, i, value):
    ...
```

For sequence-type objects, the `i` argument specifies the position in the sequence to be modified. For mappings, the `i` argument is the key under which the `value` is to be stored.

### 19.4. Static methods

A static method of a Python class is a method that does not obey the usual convention in which `self`, an instance of the class, is the first argument to the method.

To declare a static method, declare the method normally, and wrap it with the built-in `staticmethod` function, using either of the techniques described in Section 14.20, “`staticmethod()`: Create a static method” (p. 68).

Once you have declared a method to be static, the arguments you pass it are exactly the arguments it receives. Example:

```python
>>> class Hobbs:
...     @staticmethod
...     def represent():
...         print "Hobbs represent!"
...     ...

>>> Hobbs.represent()
Hobbs represent!
```
19.5. Class methods

A class method is similar to a static method (see Section 19.4, “Static methods” (p. 108)), except that its first argument is always the class that contains the method.

To declare a class method, use a declaration of this general form:

```python
@classmethod
def methodName(cls, ...):
    ...
```

When this method is called, the first argument (cls) will be the class containing `methodName`.

There are two ways to call a class method: using its class `C`, or an instance `i`. These two general forms are:

```python
C.methodName(...)  # In the first case, the class C is passed as the cls argument of the method.
i.methodName(...)  # In the second case, the class of instance i is passed as the cls argument.
```

```python
>>> class Jal(object):
    ...
    @classmethod
    ...
    def whatClass(cls, n):
    ...
        print "cls=%s n=%s" % (cls, n)
    ...
    def __init__(self, color):
    ...
        self.color = color
    ...

>>> Jal.whatClass(5)
cls=<class '__main__.Jal'> n=5
>>> eunice=Jal('green')
>>> eunice.whatClass(17)
cls=<class '__main__.Jal'> n=17
```

20. pdb: The Python interactive debugger

The Python debugger allows you to monitor and control execution of a Python program, to examine the values of variables during execution, and to examine the state of a program after abnormal termination (post-mortem analysis).

- Section 20.1, “Starting up pdb” (p. 109).
- Section 20.2, “Functions exported by pdb” (p. 110).
- Section 20.3, “Commands available in pdb” (p. 110).

20.1. Starting up pdb

To execute any Python statement under the control of pdb:

```python
>>> import pdb
>>> import yourModule
>>> pdb.run('yourModule.test()')  # Or any other statement
```

where `yourModule.py` contains the source code you want to debug.
To debug a Python script named myscript.py:

```
python /usr/lib/python2.5/pdb.py myscript.py
```

To perform post-mortem analysis:

```
>>> import pdb
>>> import yourModule
>>> yourModule.test()
[crash traceback appears here]
>>> pdb.pm()
(pdb)
```

Then you can type debugger commands at the (pdb) prompt.

### 20.2. Functions exported by pdb

The `pdb` module exports these functions:

- **`pdb.run(stmt[,globals[,locals]])`**
  
  Executes any Python statement. You must provide the statement as a string. The debugger prompts you before beginning execution. You can provide your own global name space by providing a `globals` argument; this must be a dictionary mapping global names to values. Similarly, you can provide a local namespace by passing a dictionary `locals`.

- **`pdb.runeval(expr[,globals[,locals]])`**
  
  This is similar to the `pdb run` command, but it evaluates an expression, rather than a statement. The `expr` is any Python expression in string form.

- **`pdb.runcall(func[,arg]...)`**
  
  Calls a function under `pdb` control. The `func` must be either a function or a method. Arguments after the first argument are passed to your function.

### 20.3. Commands available in pdb

The debugger prompts with a line like this:

```
(pdb)
```

At this prompt, you can type any of the `pdb` commands discussed below. You can abbreviate any command by omitting the characters in square brackets. For example, the `where` command can be abbreviated as simply `w`.

- **`expr`**
  
  Evaluate an expression `expr` and print its value.

- **`!stmt`**
  
  Execute a Python statement `stmt`. The “!” may be omitted if the statement does not resemble a `pdb` command.

- **(empty line)**
  
  If you press Enter at the (pdb) prompt, the previous command is repeated. The `list` command is an exception: an empty line entered after a `list` command shows you the next 11 lines after the ones previously listed.
args
Display the argument names and values to the currently executing function.

break [[filename:]lineno[,condition]]
The break command sets a breakpoint at some location in your program. If execution reaches a breakpoint, execution will be suspended and you will get back to the (pdb) prompt.

This form of the command sets a breakpoint at a specific line in a source file. Specify the line number within your source file as lineno; add the filename: if you are working with multiple source files, or if your source file hasn’t been loaded yet.

You can also specify a conditional breakpoint, that is, one that interrupts execution only if a given condition evaluates as true. For example, the command break 92,i>5 would break at line 92 only when i is greater than 5.

When you set a breakpoint, pdb prints a “breakpoint number.” You will need to know this number to clear the breakpoint.

break [function[,condition]]
This form of the break command sets a breakpoint on the first executable statement of the given function.

continue
Resume execution until the next breakpoint (if any).

clear [lineno]
If used without an argument, clears all breakpoints. To clear one breakpoint, give its breakpoint number (see break above).

help [cmd]
Without an argument, prints a list of valid commands. Use the cmd argument to get help on command cmd.

list [begin[,end]]
Displays your Python source code. With no arguments, it shows 11 lines centered around the current point of execution. The line about to be executed is marked with an arrow (->), and the letter B appears at the beginning of lines with breakpoints set.

To look at a given range of source lines, use the begin argument to list 11 lines around that line number, or provide the ending line number as an end argument. For example, the command list 50,65 would list lines 50-65.

next
Like step, but does not stop upon entry to a called function.

quit
Exit pdb.

return
Resume execution until the current function returns.

step
Single step: execute the current line. If any functions are called in the current line, pdb will break upon entering the function.

tbreak
Same options and behavior as break, but the breakpoint is temporary, that is, it is removed after the first time it is hit.
21. Commonly used modules

The sections below discuss only a tiny fraction of the official and unofficial module library of Python. For a full set, consult the Python Library Reference. If you want to use any of these modules, you must import them. See Section 15.6, “The import statement: Use a module” (p. 75).

21.1. math: Common mathematical operations

This module provides the usual basic transcendental mathematical functions. All trig functions use angles in radians. (For a similar set of functions using the complex number system, see the Python library documentation for the cmath module.)

The math module has two constant attributes:

<table>
<thead>
<tr>
<th>pi</th>
<th>The constant 3.14159...</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>The base of natural logarithms, 2.71828...</td>
</tr>
</tbody>
</table>

Functions in this module include:

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>acos(x)</td>
<td>Angle (in radians) whose cosine is x, that is, arccosine of x.</td>
</tr>
<tr>
<td>asin(x)</td>
<td>Arcsine of x.</td>
</tr>
<tr>
<td>atan(x)</td>
<td>Arctangent of x.</td>
</tr>
<tr>
<td>atan2(y,x)</td>
<td>Angle whose slope is y/x, even if x is zero.</td>
</tr>
<tr>
<td>ceil(x)</td>
<td>True ceiling function, defined as the smallest integer that is greater than or equal to x. For example, ceil(3.9) yields 4.0, while ceil(-3.9) yields -3.0.</td>
</tr>
<tr>
<td>cos(x)</td>
<td>Cosine of x, where x is expressed in radians.</td>
</tr>
<tr>
<td>cosh(x)</td>
<td>Hyperbolic cosine of x.</td>
</tr>
<tr>
<td>exp(x)</td>
<td>e to the x power.</td>
</tr>
<tr>
<td>floor(x)</td>
<td>True floor function, defined as the largest integer that is less than or equal to x. For example, floor(3.9) is 3.0, and floor(-3.9) is -4.0.</td>
</tr>
<tr>
<td>frexp(x)</td>
<td>Given a float x, returns a tuple (m, e) where m is the mantissa and e the binary exponent of x, and x is equal to m * (2**e). If x is zero, returns the tuple (0.0, 0).</td>
</tr>
<tr>
<td>fmod(x,y)</td>
<td>Returns (x-int(x/y))*y.</td>
</tr>
<tr>
<td>ldexp(x, i)</td>
<td>Returns x * (2**i).</td>
</tr>
<tr>
<td>hypot(x,y)</td>
<td>The square root of (x^2+y^2).</td>
</tr>
<tr>
<td>log(x)</td>
<td>Natural log of x.</td>
</tr>
<tr>
<td>log10(x)</td>
<td>Common log (base 10) of x.</td>
</tr>
</tbody>
</table>

---

15 http://docs.python.org/lib/
16 http://docs.python.org/lib/module-cmath.html
modf(x) Returns a tuple \((f, i)\) where \(f\) is the fractional part of \(x\), \(i\) is the integral part (as a float), and both have the same sign as \(x\).

sin(x) Sine of \(x\).
sinh(x) Hyperbolic sine of \(x\).
sqrt(x) Square root of \(x\).
tan(x) Tangent of \(x\).
tanh(x) Hyperbolic tangent of \(x\).

21.2. string: Utility functions for strings

Variables and functions for working with character strings.

digits
The decimal digits: "0123456789".

hexdigits
The hexadecimal digits: "0123456789abcdefABCDEF".

letters
All the letters: "abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ".

lowercase
All the lowercase letters: "abcdefghijklmnopqrstuvwxyz".

maketrans(s, t)
Builds a translation table to be used as the first argument to the .translate() string method. The arguments \(s\) and \(t\) are two strings of the same length; the result is a translation table that will convert each character of \(s\) to the corresponding character of \(t\).

```python
>>> import string
>>> bingo=string.maketrans("lLrR", "rRlL")
>>> "Cornwall Llanfair".translate(bingo)
'Colnwarr Rranfail'
>>> "Cornwall Llanfair".translate(bingo, 'ai')
'Colnwrr Rrnfl'
```

octdigits
The octal digits: "01234567".

printable
A string containing all the printable characters.

punctuation
All printable characters that are not letters or digits, to wit:

```
"!"#$%&'()*+,-./:;<=>?@[\]^_`{|}~
```

uppercase
All the uppercase characters: "ABCDEFGHIJKLMNOPQRSTUVWXYZ".

whitespace
A string containing all the characters considered to be white space, namely:

"\t\n\x0b\x0c\r "

New Mexico Tech Computer Center

Python 2.5 quick reference

113
21.3. random: Random number generation

Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number—there are only methods to produce random numbers, and a strict arithmetic procedure of course is not such a method.

-- John von Neumann

This module provides for the generation of pseudorandom numbers.

choice(L)

Returns a randomly selected element from a sequence L.

```python
>>> for i in range(9): print choice(['s', 'h', 'd', 'c']),
... s c h c s s h d
```

normalvariate(m, s)

Generate a normally distributed pseudorandom number with mean m and standard deviation s.

randint(x, y)

Returns a random integer in the closed interval [x, y]; that is, any result r will satisfy $x \leq r \leq y$.

```python
>>> for i in range(20): print randint(1,6),
... 3 4 6 3 2 1 1 2 1 2 1 2 3 3 5 6 4 4 2
```

random()

Returns a random float in the half-open interval $[0.0, 1.0)$; that is, for any result $r$, $0.0 \leq r < 1.0$.

```python
>>> for count in range(48):
... print "%.3f" % random(),
... if (count % 12) == 11: print
... 0.012 0.750 0.899 0.339 0.371 0.561 0.358 0.931 0.822 0.990 0.682 0.847
0.245 0.541 0.992 0.151 0.394 0.335 0.702 0.885 0.986 0.350 0.417 0.748
0.918 0.103 0.109 0.328 0.423 0.180 0.203 0.689 0.600 0.794 0.201 0.088
0.564 0.920 0.906 0.469 0.510 0.818 0.142 0.589 0.590 0.290 0.650 0.889
```

randrange([start,]stop[,step])

Return a random element from the sequence range(start, stop, step).

```python
>>> from random import *
>>> for i in range(35): print randrange(4),
... 0 2 2 2 1 1 2 3 1 3 3 2 2 2 3 0 2 0 0 1 2 0 2 1 1 2 2 2 1 1 3 1 1 2
>>> for i in range(35): print randrange(1,5),
... 3 3 2 1 1 1 4 4 3 2 1 1 3 2 1 2 4 1 4 2 4 4 1 1 1 1 1 4 4 1 1 2 2 1
```
>>> range(2,18,5)
[2, 7, 12, 17]
>>> for i in range(28): print randrange(2,18,5),
... 12 2 7 2 17 17 7 7 12 17 2 7 17 12 7 7 12 17 17 7 12 7 7 7 7 7 7

shuffle(L)

Randomly permute the elements of a sequence L.

Here's an example. First we build a (small) deck of cards, using a list comprehension to build a list of all possible combinations of three ranks (ace, king, queen) and four suits (spades, hearts, diamonds, and clubs). Then we shuffle the deck twice and inspect the results.

>>> ranks = 'AKQ'
>>> suits = 'shdc'
>>> deck = [ r+s
... for s in suits
... for r in ranks ]
>>> deck
['As', 'Ks', 'Qs', 'Ah', 'Kh', 'Qh', 'Ad', 'Kd', 'Qd', 'Ac', 'Kc', 'Qc']
>>> shuffle(deck)
>>> deck
['Qh', 'Ks', 'Kh', 'As', 'Kc', 'Kd', 'Qd', 'Qc', 'Ah', 'Ad', 'Qs', 'Ac']
>>> shuffle(deck)
>>> deck
['As', 'Qs', 'Ks', 'Kc', 'Ad', 'Kh', 'Qh', 'Ac', 'Ah', 'Qc', 'Qd', 'Kd']

uniform(x,y)

Returns a random float in the half-open interval \([x,y)\); that is, each result \(r\) will satisfy \(x \leq r < y\).

An assortment of other pseudorandom distributions is available. See the Python Library Reference\(^\text{18}\) for details.

### 21.4. time: Clock and calendar functions

This module provides minimal time and date functions. For a newer and more complete module, see the datetime module\(^\text{19}\).

- **Epoch time** is the time in seconds since some arbitrary starting point. For example, Unix measures time in seconds since January 1, 1970.
- **UTC** is Coordinated Universal Time, the time on the zero meridian (which goes through London).
- **DST** refers to Daylight Savings Time.

A **time tuple** is a 9-tuple \(T\) with these elements, all integers:

| \(T[0]\) | Four-digit year. | \(T[5]\) | Second, in \([0,59]\). |
| \(T[1]\) | Month, 1 for January, 12 for December. | \(T[6]\) | Day of week, 0 for Monday, 6 for Sunday. |
| \(T[2]\) | Day of month, in \([1,31]\). | \(T[7]\) | Ordinal day of the year, in \([1,366]\). |

\(^{18}\) http://docs.python.org/lib/module-random.html

\(^{19}\) http://docs.python.org/lib/module-datetime
DST flag:
1 if the time is DST, 0 if it is not DST, and -1 if unknown.


$T[8]$ DST flag: 1 if the time is DST, 0 if it is not DST, and -1 if unknown.

**Contents of the time module:**

**altzone**
The local DST offset, in seconds west of UTC (negative for east of UTC).

**asctime($T$)**
For a time-tuple $T$, returns a string of exactly 24 characters with this format:

"Thu Jun 12 15:25:31 1997"

The default time is now.

**clock()**
The accumulated CPU time of the current process in seconds, as a float.

**ctime($E$)**
Converts an epoch time $E$ to a string with the same format as `asctime()`. The default time is now.

**daylight**
Nonzero if there is a DST value defined locally.

**gmtime($E$)**
Returns the time-tuple corresponding to UTC at epoch time $E$; the DST flag will be zero. The default time is now.

**localtime($E$)**
Returns the time-tuple corresponding to local time at epoch time $E$. The default time is now.

**mktime($T$)**
Converts a time-tuple $T$ to epoch time as a float, where $T$ is the local time.

**sleep($s$)**
Suspend execution of the current process for $s$ seconds, where $s$ can be a float or integer.

**strftime($f,t$)**
Time formatting function; formats a time-tuple $t$ according to format string $f$. The default time $t$ is now. As with the string format operator, format codes start with %, and other text appears unchanged in the result. See the table of codes below.

**time()**
The current epoch time, as a float.

**timezone**
The local non-DST time zone as an offset in seconds west of UTC (negative for east of UTC). This value applies when daylight savings time is not in effect.

**tzname**
A 2-tuple $(s, d)$ where $s$ is the name of the non-DST time zone locally and $d$ is the name of the local DST time zone. For example, in Socorro, NM, you get ("MST", "MDT").

Format codes for the `strftime` function include:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%a</td>
<td>Abbreviated weekday name, e.g., 'Tue'.</td>
</tr>
<tr>
<td>%A</td>
<td>Full weekday name, e.g., 'Tuesday'.</td>
</tr>
</tbody>
</table>
Abbreviated month name, e.g., ‘Jul’.

Full month name, e.g., ‘July’.

Day of the month, two digits with left zero fill; e.g. ‘03’.

Hour on the 24-hour clock, two digits with zero fill.

Hour on the 12-hour clock, two digits with zero fill.

Day of the year as a decimal number, three digits with zero fill, e.g. ‘366’.

Decimal month, two digits with zero fill.

Minute, two digits with zero fill.

Either ‘AM’ or ‘PM’. Midnight is considered AM and noon PM.

Second, two digits with zero fill.

Numeric weekday: 0 for Sunday, 6 for Saturday.

Two-digit year. Not recommended!

Four-digit year.

If there is a time zone, a string representing that zone; e.g., ‘PST’.

Outputs the character %.

## 21.5. re: Regular expression pattern-matching

The `re` module provides functions for matching strings against regular expressions. See the O'Reilly book *Mastering Regular Expressions* by Friedl and Oram for the whys and hows of regular expressions. We discuss only the most common functions here. Refer to the *Python Library Reference* for the full feature set.

- Section 21.5.1, “Characters in regular expressions” (p. 117).
- Section 21.5.2, “Functions in the `re` module” (p. 118).
- Section 21.5.3, “Compiled regular expression objects” (p. 119).
- Section 21.5.4, “Methods on a `MatchObject`” (p. 120).

### 21.5.1. Characters in regular expressions

*Note: The raw string notation `r'...'` is most useful for regular expressions; see raw strings (p. 14), above.*

These characters have special meanings in regular expressions:

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>Matches any character except a newline.</td>
</tr>
<tr>
<td>^</td>
<td>Matches the start of the string.</td>
</tr>
<tr>
<td>$</td>
<td>Matches the end of the string.</td>
</tr>
<tr>
<td>r*</td>
<td>Matches zero or more repetitions of regular expression r.</td>
</tr>
<tr>
<td>r+</td>
<td>Matches one or more repetitions of r.</td>
</tr>
<tr>
<td>r?</td>
<td>Matches zero or one r.</td>
</tr>
</tbody>
</table>

---

20 http://en.wikipedia.org/wiki/Y2k
21 http://docs.python.org/lib/module-re.html
Non-greedy form of \( r^* \); matches as few characters as possible. The normal * operator is greedy: it matches as much text as possible.

Non-greedy form of \( r^+ \).

Non-greedy form of \( r^? \).

Matches from \( m \) to \( n \) repetitions of \( r \). For example, \( r'x{3,5}' \) matches between three and five copies of letter 'x'; \( r'(bl){4}' \) matches the string 'blblblbl'.

Non-greedy version of the previous form.

Matches one character from a set of characters. You can put all the allowable characters inside the brackets, or use \( a-b \) to mean all characters from \( a \) to \( b \) inclusive. For example, regular expression \( r'[abc]' \) will match either 'a', 'b', or 'c'. Pattern \( r'[0-9a-zA-Z]' \) will match any single letter or digit.

\[^\ldots\]  Matches expression \( r \) followed by expression \( s \).

\( rs \)  Matches either \( r \) or \( s \).

\( (r) \)  Matches \( r \) and forms it into a group that can be retrieved separately after a match; see MatchObject, below. Groups are numbered starting from 1.

\( (?:r) \)  Matches \( r \) but does not form a group for later retrieval.

\( (?P<n>r) \)  Matches \( r \) and forms it into a named group, with name \( n \), for later retrieval.

These special sequences are recognized:

\( n \)  Matches the same text as a group that matched earlier, where \( n \) is the number of that group. For example, \( r'([a-zA-Z]+):\1' \) matches the string "foo:foo".

\( A \)  Matches only at the start of the string.

\( b \)  Matches the empty string but only at the start or end of a word (where a word is set off by whitespace or a non-alphanumeric character). For example, \( r'foo\b' \) would match "foo" but not "foot".

\( B \)  Matches the empty string when not at the start or end of a word.

\( d \)  Matches any digit.

\( D \)  Matches any non-digit.

\( s \)  Matches any whitespace character.

\( S \)  Matches any non-whitespace character.

\( w \)  Matches any alphanumeric character.

\( W \)  Matches any non-alphanumeric character.

\( Z \)  Matches only at the end of the string.

\( \" \)  Matches a backslash (\) character.

21.5.2. Functions in the re module

There are two ways to match regular expressions with the re module. Assuming you import the module with import re, you can test whether a regular expression \( r \) matches a string \( s \) with the construct:

```
re.match(r,s)
```
However, if you will be matching the same regular expression many times, the performance will be better if you compile the regular expression like this:

```python
re.compile(r)
```

The `re.compile()` function returns a compiled regular expression object. You can then check a string `s` for matching by using the `.match(s)` method on that object.

Here are the functions in module `re`:

**compile(r[, f])**

Compile regular expression `r`. This function returns a compiled regular expression object; see Section 21.5.3, “Compiled regular expression objects” (p. 119). To get case-insensitive matching, use `re.I` as the `f` argument. There are other flags that may be passed to the `f` argument; see the Python Library Reference.

**match(r,s[, f])**

If `r` matches the start of string `s`, return a `MatchObject` (see below), otherwise return `None`.

**search(r,s[, f])**

Like the `match()` method, but matches `r` anywhere in `s`, not just at the beginning.

**split(r,s[, maxsplit=m])**

Splits string `s` into pieces where pattern `r` occurs. If `r` does not contain groups, returns a list of the parts of `s` that match `r`, in order. If `r` contains groups, returns a list containing all the characters from `s`, with parts matching `r` in separate elements from the non-matching parts. If the `m` argument is given, it specifies the maximum number of pieces that will be split, and the leftovers will be returned as an extra string at the end of the list.

**sub(r,R,s[, count=c])**

Replace the leftmost non-overlapping parts of `s` that match `r` using `R`; returns `s` if there is no match. The `R` argument can be a string or a function that takes one `MatchObject` argument and returns the string to be substituted. If the `c` argument is supplied (defaulting to 0), no more than `c` replacements are done, where a value of 0 means do them all.

### 21.5.3. Compiled regular expression objects

Compiled regular expression objects returned by `re.compile()` have these methods:

**.match(s,[ps][,pe])**

If the start of string `s` matches, return a `MatchObject`; if there is no match, return `None`. If `ps` is given, it specifies the index within `s` where matching is to start; this defaults to 0. If `pe` is given, it specifies the maximum length of `s` that can be used in matching.

**.pattern**

The string from which this object was compiled.

**.search(s[,ps][,pe])**

Like `match()`, but matches anywhere in `s`.

**.split(s[,maxsplit=m])**

Like `re.split()`.

**.sub(R,s[,count=c])**

Like `re.sub()`.

---

22 http://docs.python.org/lib/module-re.html
21.5.4. Methods on a MatchObject

A MatchObject is the object returned by .match() or other methods. Such an object has these methods:

.end([n])
Returns the location where a match ended. If no argument is given, returns the index of the first character past the match. If n is given, returns the index of the first character past where the nth group matched.

.endpos
The effective $p_e$ value passed to .match() or .search().

.group([n])
Retrieves the text that matched. If there are no arguments, returns the entire string that matched. To retrieve just the text that matched the nth group, pass in an integer n, where the groups are numbered starting at 1. For example, for a MatchObject $m$, $m.group(2)$ would return the text that matched the second group, or None if there were no second group.

If you have named the groups in your regular expression using a construct of the form (?P<name>...), the n argument can be the name as a string. For example, if you have a group (?P<year>\d\{4\}) (which matches four digits), you can retrieve that field using $m.group('year').$

.groups()
Return a tuple $(s_1,s_2,...)$ containing all the matched strings, where $s_i$ is the string that matched the ith group.

.pos
The effective $p_s$ value passed to .match() or .search().

.re
The regular expression object used to produce this MatchObject.

.span([n])
Returns a 2-tuple $(m.start(n),m.end(n))$.

.start([n])
Returns the location where a match started. If no argument is given, returns the index within the string where the entire match started. If an argument $n$ is given, returns the index of the start of the match for the nth group.

.string
The $s$ argument passed to .match() or .search().

21.6. sys: Universal system interface

The services in this module give you access to command line arguments, standard input and output streams, and other system-related facilities.

.argv
sys.argv[0] is the name of your Python script, or '-c' if in interactive mode. The remaining elements, sys.argv[1:], are the command line arguments, if any.

builtin_module_names
A list of the names of the modules compiled into your installation of Python.
exit(n)
   Terminate execution with status n.

modules
   A dictionary of the modules already loaded.

path
   The search path for modules, a list of strings in search order.

   Note: You can modify this list. For example, if you want Python to search directory 
   /u/dora/python/lib for modules to import before searching any other directory, these two lines will do it:

   | import sys
   | sys.path.insert(0, "/u/dora/python/lib")

platform
   A string identifying the software architecture.

stderr
   The standard error stream as a file object.

stdin
   The standard input stream as a file object.

stdout
   The standard output stream as a file object.

21.7. os: The operating system interface

   The variables and methods in the os module allow you to interact with files and directories. In most
   cases the names and functionalities are the same as the equivalent C language functions, so refer to
   Kernighan and Ritchie, The C Programming Language, second edition, or the equivalent for more details.

   chdir(p)
      Change the current working directory to that given by string p

   chmod(p,m)
      Change the permissions for pathname p to m. See module stat, below, for symbolic constants to
      be used in making up m values.

   chown(p,u,g)
      Change the owner of pathname p to user id u and group id g.

environ
   A dictionary whose keys are the names of all currently defined environmental variables, and whose
   values are the values of those variables.

error
   The exception raised for errors in this module.

execv(p,A)
   Replace the current process with a new process executing the file at pathname p, where A is a list
   of the strings to be passed to the new process as command line arguments.

execve(p,A,E)
   Like execv(), but you supply a dictionary E that defines the environmental variables for the new
   process.
_exit(n)
Exit the current process and return status code n. This method should be used only by the child process after a fork(); normally you should use sys.exit().

fork()
Fork a child process. In the child process, this function returns 0; in the parent, it returns the child’s process ID.

getcwd()
Returns the current working directory name as a string.

getegid()
Returns the effective group ID.

geteuid()
Returns the effective user ID.

getgid()
Returns the current process’s group ID. To decode user IDs, see the grp standard module\(^{23}\).

getpid()
Returns the current process’s process ID.

getppid()
Returns the parent process’s PID (process ID).

getuid()
Returns the current process’s user ID. To decode user IDs, see the pwd standard module\(^{24}\).

kill(p, s)
Send signal s to the process whose process ID is p.

link(s, d)
Create a hard link to s and call the link d.

listdir(p)
Return a list of the names of the files in the directory whose pathname is p. This list will never contain the special entries ‘.’ and ‘..’ for the current and parent directories. The entries may not be in any particular order.

lstat(p)
Like stat(), but if p is a link, you will get the status tuple for the link itself, rather than the file it points at.

mkdir(p[, m])
Create a directory at pathname p. You may optionally specify permissions m; see module stat below for the interpretation of permission values.

mkfifo(p, m)
Create a named pipe with name p and open mode m. The server side of the pipe will open it for reading, and the client side for writing. This function does not actually open the fifo, it just creates the rendezvous point.

nice(i)
Renice (change the priority) of the current process by adding i to its current priority.

\(^{23}\) http://docs.python.org/library/pwd.html
\(^{24}\) http://docs.python.org/library/pwd.html
readlink(\(p\))
   If \(p\) is the pathname to a soft (symbolic) link, this function returns the pathname to which that link points.

remove(\(p\))
   Removes the file with pathname \(p\), as in the Unix rm command. Raises OSError if it fails.

rename(\(p_o\), \(p_n\))
   Rename path \(p_o\) to \(p_n\).

rmdir(\(p\))
   Remove the directory at path \(p\).

stat(\(p\))
   Return a status tuple describing the file or directory at pathname \(p\). See module stat, below, for the interpretation of a status tuple. If \(p\) is a link, you will get the status tuple of the file to which \(p\) is linked.

symlink(\(s\), \(d\))
   Create a symbolic link to path \(s\), and call the link \(d\).

system(\(c\))
   Execute the command in string \(c\) as a sub-shell. Returns the exit status of the process.

times()
   Returns a tuple of statistics about the current process’s elapsed time. This tuple has the form \((u, s, u', s', r)\) where \(u\) is user time, \(s\) is system time, \(u'\) and \(s'\) are user and system time including all child processes, and \(r\) is elapsed real time. All values are in seconds as floats.

umask(\(m\))
   Sets the “umask” that determines the default permissions for newly created files. Returns the previous value. Each bit set in the umask corresponds to a permission that is not granted by default.

uname()
   Returns a tuple of strings describing the operating system’s version: \((s, n, r, v, m)\) where \(s\) is the name of the operating system, \(n\) is the name of the processor (node) where you are running, \(r\) is the operating system’s version number, \(v\) is the major version, and \(m\) describes the type of processor.

utime(\(p\), \(t\))
   The \(t\) argument must be a tuple \((a, m)\) where \(a\) and \(m\) are epoch times. For pathname \(p\), set the last access time to \(a\) and the last modification to \(m\).

wait()
   Wait for the termination of a child process. Returns a tuple \((p, e)\) where \(p\) is the child’s process ID and \(e\) is its exit status.

waitpid(\(p\), \(o\))
   Like wait(), but it waits for the process whose ID is \(p\). The option value \(o\) specifies what to do if the child is still running. If \(o\) is 0, you wait for the child to terminate. Use a value of os.WNOHANG if you don’t want to wait.

WNOHANG
   See waitpid() above.

### 21.8. stat: Interpretation of file status

The stat module contains a number of variables used in encoding and decoding various items returned by certain methods in the os module, such as stat() and chmod().
First, there are constants for indexing the components of a “status tuple” such as that returned by `os.stat()`:

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST_ATIME</td>
<td>The epoch time of last access (see the <code>time</code> module for interpretation of times).</td>
</tr>
<tr>
<td>ST_CTIME</td>
<td>The epoch time of the file's last status change.</td>
</tr>
<tr>
<td>ST_DEV</td>
<td>The device number.</td>
</tr>
<tr>
<td>ST_GID</td>
<td>The group ID.</td>
</tr>
<tr>
<td>ST_INO</td>
<td>The i-node number.</td>
</tr>
<tr>
<td>ST_MODE</td>
<td>The file's permissions.</td>
</tr>
<tr>
<td>ST_MTIME</td>
<td>The epoch time of last modification.</td>
</tr>
<tr>
<td>ST_NLINK</td>
<td>The number of hard links.</td>
</tr>
<tr>
<td>ST_SIZE</td>
<td>The current size in bytes.</td>
</tr>
<tr>
<td>ST_UID</td>
<td>The user ID.</td>
</tr>
</tbody>
</table>

The following functions are defined in the `stat` module for testing a mode value $m$, where $m$ is the `ST_MODE` element of the status tuple. Each function is a predicate:

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>S_ISBLK(m)</code></td>
<td>Is this a block device?</td>
</tr>
<tr>
<td><code>S_ISCHR(m)</code></td>
<td>Is this a character device?</td>
</tr>
<tr>
<td><code>S_ISDIR(m)</code></td>
<td>Is this a directory?</td>
</tr>
<tr>
<td><code>S_ISFIFO(m)</code></td>
<td>Is this a FIFO?</td>
</tr>
<tr>
<td><code>S_ISLNK(m)</code></td>
<td>Is this a soft (symbolic) link?</td>
</tr>
<tr>
<td><code>S_ISREG(m)</code></td>
<td>Is this an ordinary file?</td>
</tr>
<tr>
<td><code>S_ISSOCK(m)</code></td>
<td>Is this a socket?</td>
</tr>
</tbody>
</table>

These constants are defined for use as mask values in testing and assembling permission values such as those returned by `os.stat()` in Section 21.7, “`os`: The operating system interface” (p. 121).
21.9. os.path: File and directory interface

These functions allow you to deal with path names and directory trees. To use a given method in this module, import the os module and then use os.path.method().

For example, to get the base name of a path p, use os.path.basename(p).

abs(path(p))
Return the absolute path name that is equivalent to path p.

basename(path(p))
Return the base name portion of a path name string p. See split(), below.

commonprefix(list)  
For a list L containing pathname strings, return the longest string that is a prefix of each element in L.

exists(path(p))
Predicate for testing whether pathname p exists.

expanduser(path(p))
If p is a pathname starting with a tilde character (~), return the equivalent full pathname; otherwise return p.

isabs(path(p))
Predicate for testing whether p is an absolute pathname (e.g., starts with a slash on Unix systems).

isfile(path(p))
Predicate for testing whether p refers to a regular file, as opposed to a directory, link, or device.

islink(path(p))
Predicate for testing whether p is a soft (symbolic) link.

ismount(path(p))
Predicate for testing whether p is a mount point, that is, whether p is on a different device than its parent directory.

join(path(p), path(q))  
If q is an absolute path, returns q. Otherwise, if p is empty or ends in a slash, returns p+q, but otherwise it returns p+'/'+q.

normcase(path(p))
Return pathname p with its case normalized. On Unix systems, this does nothing, but on Macs it lowercases p.

samefile(path(p), path(q))
Predicate for testing whether p and q are the same file (that is, the same inode on the same device).  
This method may raise an exception if os.stat() fails for either argument.

split(path(p))
Return a 2-tuple (H, T) where T is the tail end of the pathname (not containing a slash) and H is everything up to the tail. If p ends with a slash, returns (p, '')  
If p contains no slashes, returns ('', p). The returned H string will have its trailing slash removed unless H is the root directory.

splitext(path(p))
Returns a 2-tuple (R, E) where E is the “extension” part of the pathname and R is the “root” part.  
If p contains at least one period, E will contain the last period and everything after that, and R will be everything up to but not including the last period. If p contains no periods, returns ('', p).
**walk(p,V,a)**

Walks an entire directory structure starting at pathname `p`. See below for more information.

The `os.path.walk(p,V,a)` function does the following for every directory at or below `p` (including `p` if `p` is a directory), this method calls the “visitor function” `V` with arguments

<table>
<thead>
<tr>
<th>V(a,d,N)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>a</code></td>
</tr>
<tr>
<td><code>d</code></td>
</tr>
<tr>
<td><code>N</code></td>
</tr>
</tbody>
</table>