This is a Jupyter Notebook adaptation of the Python 3 tutorial[3].

1 Data Structures

This chapter describes some things you’ve learned about already in more detail, and adds some new things as well.

1.1 More on Lists

The list data type has some more methods. Here are all of the methods of list objects:

- `list.append(x)`
  
  Add an item to the end of the list. Equivalent to `a[len(a):] = [x]`.

- `list.extend(iterable)`
  
  Extend the list by appending all the items from the iterable. Equivalent to `a[len(a):] = iterable`.

- `list.insert(i, x)`
  
  Insert an item at a given position. The first argument is the index of the element before which to insert, so `a.insert(0, x)` inserts at the front of the list, and `a.insert(len(a), x)` is equivalent to `a.append(x)`.

- `list.remove(x)`
  
  Remove the first item from the list whose value is `x`. It is an error if there is no such item.

- `list.pop([i])`
  
  Remove the item at the given position in the list, and return it. If no index is specified, `a.pop()` removes and returns the last item in the list. (The square brackets around the `i` in the method signature denote that the parameter is optional, not that you should type square brackets at that position. You will see this notation frequently in the Python Library Reference)*

- `list.clear()`
  
  Remove all items from the list. Equivalent to `del a[:]`.

- `list.index(x[, start[, end]])`
  
  Return zero-based index in the list of the first item whose value is `x`. Raises a `ValueError` if there is no such item.

  The optional arguments `start` and `end` are interpreted as in the slice notation and are used to limit the search to a particular subsequence of the list. The returned index is computed relative to the beginning of the full sequence rather than the `start` argument.*

- `list.count(x)`
  
  Return the number of times `x` appears in the list.*

- `list.sort(key=None, reverse=False)`
Sort the items of the list in place (the arguments can be used for sort customization, see sorted for their explanation).

- `list.reverse()`
  Reverse the elements of the list in place.
- `list.copy()`
  Return a shallow copy of the list. Equivalent to `a[:].`

An example that uses most of the list methods:

In [ ]:
```python
fruits = ['orange', 'apple', 'pear', 'banana', 'kiwi', 'apple', 'banana']
fruits.count('apple')
```

In [ ]:
```python
fruits.count('tangerine')
```

In [ ]:
```python
fruits.index('banana')
```

In [ ]:
```python
fruits.index('banana', 4)  # Find next banana starting a position 4
```

In [ ]:
```python
fruits.reverse()
fruits
```

In [ ]:
```python
fruits.append('grape')
fruits
```

In [ ]:
```python
fruits.sort()
fruits
```

In [ ]:
```python
fruits.pop()
```

You might have noticed that methods like `insert`, `remove` or `sort` that only modify the list have no return value printed -- they return the default `None`. This is a design principle for all mutable data structures in Python.

### 1.1.1 Using Lists as Stacks

The list methods make it very easy to use a list as a stack, where the last element added is the first element retrieved ("last-in, first-out"). To add an item to the top of the stack, use `append`. To retrieve an item from the top of the stack, use `pop` without an explicit index. For example:

In [ ]:
```python
stack = [3, 4, 5]
stack.append(6)
stack.append(7)
stack
```

In [ ]:
```python
stack.pop()
```

In [ ]:
```python
stack
```

In [ ]:
```python
stack.pop()
```

In [ ]:
```python
stack.pop()
```

In [ ]:
```python
stack
```
1.1.2 Using Lists as Queues

It is also possible to use a list as a queue, where the first element added is the first element retrieved ("first-in, first-out"); however, lists are not efficient for this purpose. While appends and pops from the end of list are fast, doing inserts or pops from the beginning of a list is slow (because all of the other elements have to be shifted by one).

To implement a queue, use collections.deque which was designed to have fast appends and pops from both ends. For example:

```
In [ ]: from collections import deque
    queue = deque(["Eric", "John", "Michael"])
    queue.append("Terry")  # Terry arrives
    queue.append("Graham")  # Graham arrives

In [ ]: queue.popleft()  # The first to arrive now leaves

In [ ]: queue.popleft()  # The second to arrive now leaves

In [ ]: queue  # Remaining queue in order of arrival
```

1.1.3 List Comprehensions

List comprehensions provide a concise way to create lists. Common applications are to make new lists where each element is the result of some operations applied to each member of another sequence or iterable, or to create a subsequence of those elements that satisfy a certain condition.

For example, assume we want to create a list of squares, like:

```
In [ ]: squares = []
    for x in range(10):
       squares.append(x**2)

In [ ]: squares
```

Note that this creates (or overwrites) a variable named x that still exists after the loop completes. We can calculate the list of squares without any side effects using:

```
In [ ]: squares = list(map(lambda x: x**2, range(10)))

or, equivalently:

In [ ]: squares = [x**2 for x in range(10)]
```

which is more concise and readable.

A list comprehension consists of brackets containing an expression followed by a for clause, then zero or more for or if clauses. The result will be a new list resulting from evaluating the expression in the context of the for and if clauses which follow it. For example, this listcomp combines the elements of two lists if they are not equal:

```
In [ ]: [(x, y) for x in [1,2,3] for y in [3,1,4] if x != y]
```
and it’s equivalent to:

```python
In []: combs = []
    for x in [1,2,3]:
        for y in [3,1,4]:
            if x != y:
                combs.append((x, y))

In []: combs
```

Note how the order of the `for` and `if` statements is the same in both these snippets. If the expression is a tuple (e.g. the `(x, y)` in the previous example), it must be parenthesized.

```python
In []: vec = [-4, -2, 0, 2, 4]
    # create a new list with the values doubled
    [x*2 for x in vec]

In []: # filter the list to exclude negative numbers
    [x for x in vec if x >= 0]

In []: # apply a function to all the elements
    [abs(x) for x in vec]

In []: # call a method on each element
    freshfruit = ['banana', 'loganberry', 'passion fruit']
    [weapon.strip() for weapon in freshfruit]

In []: # create a list of 2-tuples like (number, square)
    [(x, x**2) for x in range(6)]

In []: # the tuple must be parenthesized, otherwise an error is raised
    [x, x**2 for x in range(6)]

In []: # flatten a list using a listcomp with two 'for'
    vec = [[1,2,3], [4,5,6], [7,8,9]]
    [num for elem in vec for num in elem]
```

List comprehensions can contain complex expressions and nested functions:

```python
In []: from math import pi
    [str(round(pi, i)) for i in range(1, 6)]
```

1.1.4 Nested List Comprehensions

The initial expression in a list comprehension can be any arbitrary expression, including another list comprehension.

Consider the following example of a 3x4 matrix implemented as a list of 3 lists of length 4:
In [ ]: matrix = [
    [1, 2, 3, 4],
    [5, 6, 7, 8],
    [9, 10, 11, 12],
]

The following list comprehension will transpose rows and columns:

In [ ]: [[row[i] for row in matrix] for i in range(4)]

As we saw in the previous section, the nested listcomp is evaluated in the context of the for that follows it, so this example is equivalent to:

In [ ]: transposed = []
    for i in range(4):
        transposed.append([row[i] for row in matrix])

In [ ]: transposed

which, in turn, is the same as:

In [ ]: transposed = []
    for i in range(4):
        # the following 3 lines implement the nested listcomp
        transposed_row = []
        for row in matrix:
            transposed_row.append(row[i])
        transposed.append(transposed_row)

In [ ]: transposed

In the real world, you should prefer built-in functions to complex flow statements. The zip function would do a great job for this use case:

In [ ]: list(zip(*matrix))

1.2 The del statement

There is a way to remove an item from a list given its index instead of its value: the del statement. This differs from the pop method which returns a value. The del statement can also be used to remove slices from a list or clear the entire list (which we did earlier by assignment of an empty list to the slice). For example:

In [ ]: a = [-1, 1, 66.25, 333, 333, 1234.5]
    del a[0]

In [ ]: a

In [ ]: del a[2:4]

a
In [ ]: `del a[:]`

    a

del can also be used to delete entire variables:

In [ ]: `del a`

Referencing the name `a` hereafter is an error (at least until another value is assigned to it). We’ll find other uses for `del` later.

1.3 Tuples and Sequences

We saw that lists and strings have many common properties, such as indexing and slicing operations. They are two examples of sequence data types (see typesseq). Since Python is an evolving language, other sequence data types may be added. There is also another standard sequence data type: the tuple.

A tuple consists of a number of values separated by commas, for instance:

In [ ]: `t = 12345, 54321, 'hello!'`

    t[0]

In [ ]: `t`

In [ ]: `# Tuples may be nested:`

    u = t, (1, 2, 3, 4, 5)

    u

In [ ]: `# Tuples are immutable:`

    t[0] = 88888

In [ ]: `# but they can contain mutable objects:`

    v = ([1, 2, 3], [3, 2, 1])

    v

As you see, on output tuples are always enclosed in parentheses, so that nested tuples are interpreted correctly; they may be input with or without surrounding parentheses, although often parentheses are necessary anyway (if the tuple is part of a larger expression). It is not possible to assign to the individual items of a tuple, however it is possible to create tuples which contain mutable objects, such as lists.

Though tuples may seem similar to lists, they are often used in different situations and for different purposes. Tuples are immutable, and usually contain a heterogeneous sequence of elements that are accessed via unpacking (see later in this section) or indexing (or even by attribute in the case of namedtuples). Lists are mutable, and their elements are usually homogeneous and are accessed by iterating over the list.

A special problem is the construction of tuples containing 0 or 1 items: the syntax has some extra quirks to accommodate these. Empty tuples are constructed by an empty pair of parentheses; a tuple with one item is constructed by following a value with a comma (it is not sufficient to enclose a single value in parentheses). Ugly, but effective. For example:
In [ ]: empty = ()
    singleton = 'hello',  # <-- note trailing comma
    len(empty)

In [ ]: len(singleton)

In [ ]: singleton

The statement `t = 12345, 54321, 'hello!'` is an example of *tuple packing*; the values 12345, 54321 and 'hello!' are packed together in a tuple. The reverse operation is also possible:

In [ ]: x, y, z = t

This is called, appropriately enough, *sequence unpacking* and works for any sequence on the right-hand side. Sequence unpacking requires that there are as many variables on the left side of the equals sign as there are elements in the sequence. Note that multiple assignment is really just a combination of tuple packing and sequence unpacking.

### 1.4 Sets

Python also includes a data type for *sets*. A set is an unordered collection with no duplicate elements. Basic uses include membership testing and eliminating duplicate entries. Set objects also support mathematical operations like union, intersection, difference, and symmetric difference.

Curly braces or the `set` function can be used to create sets. Note: to create an empty set you have to use `set()`, not `{}`; the latter creates an empty dictionary, a data structure that we discuss in the next section.

Here is a brief demonstration:

In [ ]: basket = {'apple', 'orange', 'apple', 'pear', 'orange', 'banana'}
    print(basket)  # show that duplicates have been removed

In [ ]: 'orange' in basket  # fast membership testing

In [ ]: 'crabgrass' in basket

In [ ]: # Demonstrate set operations on unique letters from two words
    a = set('abracadabra')
    b = set('alacazam')
    a  # unique letters in a

In [ ]: a - b  # letters in a but not in b

In [ ]: a | b  # letters in either a or b

In [ ]: a & b  # letters in both a and b

In [ ]: a ^ b  # letters in a or b but not both

Similarly to list comprehensions, set comprehensions are also supported:

In [ ]: a = {x for x in 'abracadabra' if x not in 'abc'}
    a
1.5 Dictionaries

Another useful data type built into Python is the dictionary (see typesmapping). Dictionaries are sometimes found in other languages as "associative memories" or "associative arrays". Unlike sequences, which are indexed by a range of numbers, dictionaries are indexed by keys, which can be any immutable type; strings and numbers can always be keys. Tuples can be used as keys if they contain only strings, numbers, or tuples; if a tuple contains any mutable object either directly or indirectly, it cannot be used as a key. You can't use lists as keys, since lists can be modified in place using index assignments, slice assignments, or methods like append and extend.

It is best to think of a dictionary as an unordered set of key: value pairs, with the requirement that the keys are unique (within one dictionary). A pair of braces creates an empty dictionary: {}. Placing a comma-separated list of key:value pairs within the braces adds initial key:value pairs to the dictionary; this is also the way dictionaries are written on output.

The main operations on a dictionary are storing a value with some key and extracting the value given the key. It is also possible to delete a key:value pair with del. If you store using a key that is already in use, the old value associated with that key is forgotten. It is an error to extract a value using a non-existent key.

Performing list(d.keys()) on a dictionary returns a list of all the keys used in the dictionary, in arbitrary order (if you want it sorted, just use sorted(d.keys()) instead). To check whether a single key is in the dictionary, use the in keyword.

Here is a small example using a dictionary:

In [ ]: tel = {'jack': 4098, 'sape': 4139}
     tel['guido'] = 4127
     tel

In [ ]: tel['jack']

In [ ]: del tel['sape']
     tel['irv'] = 4127
     tel

In [ ]: list(tel.keys())

In [ ]: sorted(tel.keys())

In [ ]: 'guido' in tel

In [ ]: 'jack' not in tel

The dict constructor builds dictionaries directly from sequences of key-value pairs:

In [ ]: dict([(sape, 4139), (guido, 4127), (jack, 4098)])

In addition, dict comprehensions can be used to create dictionaries from arbitrary key and value expressions:

In [ ]: {x: x**2 for x in (2, 4, 6)}

When the keys are simple strings, it is sometimes easier to specify pairs using keyword arguments:

In [ ]: dict(sape=4139, guido=4127, jack=4098)
1.6 Looping Techniques

When looping through dictionaries, the key and corresponding value can be retrieved at the same time using the `items` method:

```python
In [ ]: knights = {'gallahad': 'the pure', 'robin': 'the brave'}
   for k, v in knights.items():
       print(k, v)
```

When looping through a sequence, the position index and corresponding value can be retrieved at the same time using the `enumerate` function:

```python
In [ ]: for i, v in enumerate(['tic', 'tac', 'toe']):
    print(i, v)
```

To loop over two or more sequences at the same time, the entries can be paired with the `zip` function:

```python
In [ ]: questions = ['name', 'quest', 'favorite color']
   answers = ['lancelot', 'the holy grail', 'blue']
   for q, a in zip(questions, answers):
       print('What is your {0}? It is {1}.'.format(q, a))
```

To loop over a sequence in reverse, first specify the sequence in a forward direction and then call the `reversed` function:

```python
In [ ]: for i in reversed(range(1, 10, 2)):
    print(i)
```

To loop over a sequence in sorted order, use the `sorted` function which returns a new sorted list while leaving the source unaltered:

```python
In [ ]: basket = ['apple', 'orange', 'apple', 'pear', 'orange', 'banana']
   for f in sorted(set(basket)):
       print(f)
```

It is sometimes tempting to change a list while you are looping over it; however, it is often simpler and safer to create a new list instead:

```python
In [ ]: import math
   raw_data = [56.2, float('NaN'), 51.7, 55.3, 52.5, float('NaN'), 47.8]
   filtered_data = []
   for value in raw_data:
       if not math.isnan(value):
           filtered_data.append(value)

In [ ]: filtered_data
```
1.7 More on Conditions

The conditions used in while and if statements can contain any operators, not just comparisons.

The comparison operators in and not in check whether a value occurs (does not occur) in a sequence. The operators is and is not compare whether two objects are really the same object; this only matters for mutable objects like lists. All comparison operators have the same priority, which is lower than that of all numerical operators.

Comparisons can be chained. For example, \( a < b == c \) tests whether \( a \) is less than \( b \) and moreover \( b \) equals \( c \).

Comparisons may be combined using the Boolean operators and and or, and the outcome of a comparison (or of any other Boolean expression) may be negated with not. These have lower priorities than comparison operators; between them, not has the highest priority and or the lowest, so that \( A \) and \( not B \) or \( C \) is equivalent to \( (A \ and \ (not \ B)) \) or \( C \). As always, parentheses can be used to express the desired composition.

The Boolean operators and and or are so-called short-circuit operators: their arguments are evaluated from left to right, and evaluation stops as soon as the outcome is determined. For example, if \( A \) and \( C \) are true but \( B \) is false, \( A \ and \ B \ and \ C \) does not evaluate the expression \( C \). When used as a general value and not as a Boolean, the return value of a short-circuit operator is the last evaluated argument.

It is possible to assign the result of a comparison or other Boolean expression to a variable. For example,

```
In [ ]: string1, string2, string3 = '', 'Trondheim', 'Hammer Dance'
    non_null = string1 or string2 or string3
    non_null
```

Note that in Python, unlike C, assignment cannot occur inside expressions. C programmers may grumble about this, but it avoids a common class of problems encountered in C programs: typing `=` in an expression when `==` was intended.

1.8 Comparing Sequences and Other Types

Sequence objects may be compared to other objects with the same sequence type. The comparison uses lexicographical ordering: first the first two items are compared, and if they differ this determines the outcome of the comparison; if they are equal, the next two items are compared, and so on, until either sequence is exhausted. If two items to be compared are themselves sequences of the same type, the lexicographical comparison is carried out recursively. If all items of two sequences compare equal, the sequences are considered equal. If one sequence is an initial subsequence of the other, the shorter sequence is the smaller (lesser) one. Lexicographical ordering for strings uses the Unicode code point number to order individual characters. Some examples of comparisons between sequences of the same type:

\[
\begin{align*}
(1, 2, 3) & \quad < (1, 2, 4) \\
[1, 2, 3] & \quad < [1, 2, 4] \\
'ABC' & < 'C' < 'Pascal' < 'Python' \\
(1, 2, 3, 4) & \quad < (1, 2, 4) \\
(1, 2) & \quad < (1, 2, -1) \\
(1, 2, 3) & \quad == (1.0, 2.0, 3.0) \\
(1, 2, ('aa', 'ab')) & \quad < (1, 2, ('abc', 'a'), 4)
\end{align*}
\]
Note that comparing objects of different types with < or > is legal provided that the objects have appropriate comparison methods. For example, mixed numeric types are compared according to their numeric value, so 0 equals 0.0, etc. Otherwise, rather than providing an arbitrary ordering, the interpreter will raise a TypeError exception.

Footnotes