(U) `__init__.py` can contain any code, but it's best to keep it short and focused on just what's needed to initialize and manage the package. For example:

- (U) setting the `__all__` variable to tell Python what modules to include when someone runs `from package import *`
- (U) automatically import some of the submodules so that when someone runs `import package`, then they can run `package.function` rather than `package.submodule.function`

(U) Installing packages

(U) Packages are actually the common way to share and distribute modules. A package can contain a single module -- there is no requirement for it to hold multiple modules. If you're wanting to work with a Python module that is not in the standard library (i.e. not installed with Python by default), then you will probably need to install the package that contains it. Python developers don't usually share or install individual module files.

(U) pip and PyPI

(U) On the command line, the standard tool for installing a package is `pip`, Python's package manager. (`pip` ships with Python by default nowadays, but if you're using an older version, you may have to install it yourself.) To use `pip`, you need to configure it to point at a package repository. On the outside, the big repository everyone uses is called PyPI (a.k.a. the Cheese Shop).

(U//~~FOUO~~) REPOMAN and nsa-pip

(U//~~FOUO~~) REPOMAN also imports and hosts a mirror of PyPI on the high side. Additionally, there is a nsa-pip server that connects to both REPOMAN's PyPI mirror and a variety of internal NSA-developed packages hosted on GitLab.

- (U//~~FOUO~~) List of internal NSA packages
- (U//~~FOUO~~) Links to some NSA package docs

(U) ipydeps & pypki2

(U//~~FOUO~~) If you are working in a Jupyter notebook, it can be awkward trying to install packages from the command line with `pip` and then use them. Instead, `ipydeps` is a module that allows you to install packages directly from the notebook. It also uses the `pypki2` module behind the scenes to handle HTTPS connections that need your PKI certificates.

```
import ipydeps
ipydeps.pip('prettytable')
```

(b)(3)-P.L. 86-36

(U//~~FOUO~~) Another thing that `ipydeps` does behind the scenes is try to install operating system (non-Python) dependencies that the package needs in order to install and run correctly. That is manually configured by the Jupyter team here at NSA. If you run into trouble installing a package with `ipydeps` in Jupyter on LABBENCH, contact [redacted] and provide the name of the package you are trying to install and the errors you are seeing.

Modules and Packages

(b)(3)-P.L. 86-36

0

Updated almost 3 years ago by [redacted] in COMP 3321

3 307 60

fcs6 access gitlab python

(U) Lesson 08: Modules and Packages

Recommendations

(U) I see you like Python, so I put Python in your Python

(U) We've seen how to write scripts; now we want to reuse the good parts. We've already used the `import` command, which lets us piggyback on the work of others–either through Python's extensive standard library or through additional, separately-installed packages. It can also be used to proactively leverage the long tail of our own personal production. In this lesson, we cover, in much greater depth, the mechanics and principles of writing and distributing modules and packages. Suppose you have a script named `my_funcs.py` in your current directory. Then the following works just fine:

```
import my_funcs
```

```
import my_funcs as m
```

```
import importlib
importlib.reload(m)
```

```
from my_funcs import string_appender
```

```
from my_funcs import * # BE CAREFUL
```

(U) If you change the source file `my_funcs.py` in between `import` commands, you will have different versions of the functions imported. So what's going on?

(U) Namespaces

(U) When you `import` a **module** (what we used to call merely a *script*), Python executes it as if from the command line, then places variables and functions inside a **namespace** defined by the script name (or using the optional `as` keyword). When you `from <module> import <name>`, the variables are imported into your current namespace. Think of a namespace as a super-variable that contains references to lots of other variables, or as a super-class that can contain data, functions, and classes.

(U) After import, a module is dynamic like any Python object; for example, the `reload` function takes a module as an argument, and you can add data and methods to the module after you've imported it (but they won't persist beyond the lifetime of your script or session).

```
import my_funcs as m
```

```
def silly_func(x):
    return "Silly {}!".format(x)
```

```
m.silly_func = silly_func
```

```
m.silly_func("Mark")
```

```
Silly Mark!
```

(U) In contrast, the `from <module> import <function>` command adds the function to the current namespace.

(U) Preventing Excess Output: The Magic of `__main__`

(U) Suppose you have a script that does something awesome, called `awesome.py`:

```
class Awesome(object):
    def __init__(self, awesome_thing):
        self.thing = awesome_thing
    def __str__(self):
        return "{0.thing} is AWESOME.".format(self)

a = Awesome("BASE Jumping")
print(a)
```

P.L. 86-36

(U) This can be executed from the command line or imported:

```
(VENV)[redacted]$ python awesome.py
BASE Jumping is AWESOME
(VENV)[redacted]$ python
```

```
import awesome
```

```
BASE Jumping is AWESOME.
```

```
a
```

```
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
NameError: name 'a' is not defined
```

```
awesome.a
```

```
<awesome.Awesome object at 0x7fa222a8b410>
```

```
print(awesome.a)
```

```
BASE Jumping is AWESOME.
```

(U) You don't want that print statement to execute every time you import it. Of equal importance, `awesome.a` is probably extraneous within an import. Let's fix it to get rid of those when you import the module, but keep them when you execute the script.

```
class Awesome(object):
    def __init__(self, awesome_thing):
        self.thing = awesome_thing
    def __str__(self):
        return "{0.thing} is AWESOME.".format(self)

if __name__ == '__main__':
    a = Awesome("BASE Jumping")
        print(a)
```

(U) We can do even better. There are some situations, e.g. profiling or testing, where we would want to import the module, then look at what would happen if we run it as a script. To enable that, move the main functionality into a function called `main()` :

```
class Awesome(object):
    def __init__(self, awesome_thing):
        self.thing = awesome_thing
    def __str__(self):
        return "{0.thing} is AWESOME.".format(self)

def main():
    a = Awesome("BASE Jumping")
    print(a)

if __name__ == '__main__':
    main()
```

(U) From Modules to Packages

(U) A single Python **module** corresponds to a **file**. It's not hard to imagine a situation where you have several related modules that you want to group together in the equivalent of a **folder**; the Python term for this concept is a **package**. We make a package by

- creating a folder
- putting scripts/modules inside it
- adding some Python Magic (which obviously will involve `__` in some way, shape, or form)

(U) For example, we'll put `awesome.py` in a package called feelings–later on, we'll add `terrible.py` and `totally_rad.py` The directory structure is:

```
feelings/
├── awesome.py
├── __init__.py
└── __main__.py
```

(U) The `__init__.py` file is **REQUIRED**; without it, Python won't identify this folder as a package. However, `__main__.py` is optional but nice; if you have it, you can type `python feelings` and the contents of `__main__.py` will be executed as a script. (NB: Now you can postulate on what `if __name__ == '__main__':` is really doing.

(U) The `__init__.py` file **can contain commands**. Much like the `__init__()` function of a `class` , the `__init__.py` is executed immediately after importing the package. One common use is to expose modules as package attributes; all this takes is `import <module_name>` in the package's `__init__.py` file.

(U) Onward to the Whole World

Pretty soon, you'll want to share the **feelings** packages with a wider audience. There are thousands of people who want to do `Awesome` stuff, but don't have the time to make their own version, which wouldn't be as good as yours anyway, so they're counting on you to provide this package in a convenient, easy-to-install manner.

(U) Shareable Packages

(U) The **_setuptools** package (which is built on **distutils**), used in conjunction with virtual environments and publicly accessible repositories in revision control systems make sharing your work as easy as `pip install` ing a package from PyPI. You are using a revision control system, aren't you? This lesson assumes that you use **git** and push your repositories to GitLab

(U) To make the **feelings** package available to the whole world, it should be placed at the root of a git repository, alongside a setup script called `setup.py` , i.e.

```
feelings_repo
├── feelings/
│   ├── awesome.py
│   ├── __init__.py
│   └── __main__.py
└── setup.py
```

(U) The `setup.py` script imports from one of two packages that handle management and installation of other packages. We'll use **setuptools** in this example, because it is more powerful and installed by default in virtual environments. In simple cases like this one, the built-in **distutils** module is more than adequate. and functionally identical.

(U) The script calls a single function, `setup` , and takes metadata about the package, including the name and version number of the package, the name and email of the developer responsible for the package, and a list of packages (or modules). It looks like this:

```
from setuptools import setup

setup(name="pyTest",
      version='0.0.1',
      description="The simplest Python Package imaginable",
      author=[redacted]
      author_email="[redacted]",
      packages=['feelings'],
      )
```

P.L. 86-36

(U) To use **distutils** instead of **setuptools**, change the first line to read `from distutils.core import setup`. Two powerful advantages of **setuptools** over **distutils** are:

- Dependency management, so that external packages available in PyPI will be installed automatically, and
- Automatic creation of *entry point* shell scripts that hook into specified functions in your code.

P.L. 86-36

(U) Sharing Packages

(U) We have bigger fish to fry–we want to get the `Awesome`-ness out into the world, and we're almost there. Once the changes have been commited and pushed to GitLab, we can share them with one simple pip command. Inside of a virtual environment, anyone with access to GitLab can execute

```
$ pip install -e git+git@gitlab.coi.nsa.ic.gov:[redacted]/feelings.git#egg=feelings
```

(U) The `-e` flag installs the repository as *editable*, a.k.a. in *developer mode*. This means that the full git repository is cloned inside the virtual environment's `src` folder and can be modified or updated in place (e.g. using `git clone`) without requiring reinstallation. The `#egg=feelings` is necessary for `pip install` to work, and must be added manually; it is neither required nor even used by GitLab.

(U) Once your user has `pip install`-ed your package, that's it! She can now do awesome stuff, like

```
from feelings import awesome

a = awesome.Awesome("Dostoyevsky")

print(a)
```

```
Dostoyevsky is AWESOME.
```

(U) Even better, it only takes little more work for her to include your package as a dependency in her packages and applications!

Lesson 09: Exceptions, Profiling, and Testing

0

(b)(3)-P.L. 86-36

Updated 8 months ago by [redacted] in COMP 3321

3 539 281

python fcs6

(U) Exception handling and code testing and profiling in Python.

Recommendations

UNCLASSIFIED

(U) Introduction

(U) Attention to exception handling, profiling, and testing distinguishes professional developers writing high-quality code from amateurs that hack around just enough to get the job done. Each topic warrants many hours of discussion on its own, but Python makes it possible to start learning and using these principles with minimal effort. This section covers basic ideas to get you interested and see the usefulness of these ideas and modules. Let's begin...by making some errors.

(U) Exceptions

(U) Python is very flexible and will try its absolute best to do whatever you ask it to, but sometimes you can just confuse it way too much. The first type of **error** is the **syntax error**. By this point in the course, we've all seen more than enough of these! They happens when Python cannot parse what you typed.

```
for i in range(10)
```

```
def altered_cool():
    print(awesome.Awesome('Artisanal vinegar'))  # still there?
    print(coolgroup('the intelligentsia'))
    cool = 'hipster'
    lumberjack = True
    print(sorted(locals().keys()))
    print(locals()['cool'])
```

```
altered_cool()
```

```
'lumberjack' in globals()
```

```
globals()['cool']
```

```
globals() == locals()
```

(U) Module namespaces

(U) Finally, each module also has its own *module* namespace. You can inspect them using the module's special `__dict__` method.

```
sorted(awesome.__dict__.keys())
```

```
# guess what?
dir(awesome) == sorted(awesome.__dict__.keys())
```

```
awesome.__dict__['cool']  # can also print this to get the memory location
```

```
# didn't we just see that here?
globals()['coolgroup']
```

```
dir(awe)
```

```
awe.__dict__['cool']
```

```
awe == awesome
```

```
id(awe) == id(awesome)
```

(U) Modifying module namespaces

(U) You can add to module namespaces on the fly. Keep in mind, though, that this will only last until the program exits, and the actual module file will be unchanged.

```
def more_awesome():
  return "They're awesome!"

awe.exclaim = more_awesome
```

```
awe.exclaim()
```

```
'exclaim' in dir(awe)
```

```
'exclaim' in dir(awesome)
```

(U) Packages

(U) What if you want to organize your code into multiple modules? Since a module is a file, the natural thing to do is to gather all your related modules into a single directory or folder. And, indeed, a Python **package** is just that: a directory that contains modules, a special `__init__.py` file, and sometimes more packages or other helper files.

```
  File "<ipython-input-1-6f7914dd2e9a>", line 1
    for i in range(10)
                      ^
SyntaxError: invalid syntax
```

(U) Python could not parse what we were trying to do here (because we forgot our colon). It did, however, let us know where things stopped making sense. Note the printed line with an tiny arrow (^) pointing to where Python thinks there is an issue.

(U) The statement `SyntaxError: invalid syntax` is an example of a special exception called a **SyntaxError**. It is fairly easy to see what happened here, and there is not much to do besides fixing your typo. Other **exceptions** can be much more interesting.

(U) There are many types of exceptions:

```
import builtins
```

```
# This will display a lot of output.
# To make it scrollable, select this cell and choose
# Cell > Current Output > Toggle Scrolling
help(builtins)
```

```
# Python 2 used to have this info in the `exceptions` module
# Python 3 moved it into `builtins` for consistency
# So for python 2, try this instead:
import exceptions
dir(exceptions)
```

(U) I bet we can make some of these happen. In fact, you probably already have recently.

```
1/0
```

```
def f():
    1/0
```

```
f()
```

```
1/'0'
```

```
import chris
```

```
file = open('data','w')
```

```
file.read()
```

(U) Exception Handling

(U) When exceptions might occur, the best course of action to is to **handle** them and do something more useful than exit and print something to the screen. In fact, sometimes exceptions can be very useful tools (e.g. `KeyboardInterrupt`). In Python, we handle exeptions with the `try` and `except` commands.

(U) Here is how it works:

1. (U) Everything between the `try` and `except` commands is executed.
2. (U) If that produces no exception, the `except` block is skipped and the program continues.
3. (U) If an exception occurs, the rest of the `try` block is skipped.
4. (U) If the type of exception is named after the `except` keyword, the code after the except command is executed.
5. (U) Otherwise, the execution stops and you have an **unhandled exception**.

(U) Everything makes more sense with an example:

```
def f(x):
    try:
        print ("I am going to convert the input to an integer")
        print (int(x))
    except ValueError:
        print ("Sorry, I was not able to convert that.")
```

```
f(2)
```

```
f('2')
```

```
f('two')
```

(U) You can add multiple Exception types to the except command:

```
... except (TypeError, ValueError):
```

(U) The keyword `as` lets us grab the message from the error:

```
def be_careful(a, b):
    try:
        print(float(a)/float(b))
    except (ValueError, TypeError, ZeroDivisionError) as detail:
        print("Handled Exception: ", detail)
    except:
        print("Unexpected error!")
    finally:
        print("THIS WILL ALWAYS RUN!")
```

```
be_careful(1,0)
```

```
be_careful(1,[1,2])
```

```
be_careful(1,'two')
```

```
be_careful(16**400,1)
```

```
float(16**400)
```

(U) We've also added the `finally` command. It will always be executed, regardless of whether there was an exception or not, so it should be used as a place to clean up anything left over from the `try` and `except` clauses, e.g. closing files that might still be open.

(U) Raising Exceptions

(U) Sometimes, you will want to cause an exception and let someone else handle it. This can be done with the `raise` command.

```
raise TypeError('You submitted the wrong type')
```

(U) If no built-in exception is suitable for what you want to raise, defining a new type of exception is as easy as creating a new class that inherits from the `Exception` type.

```
class MyPersonalError(Exception):
    pass
```

```
raise MyPersonalError("I am mighty. Hear my roar!")
```

```
def locater (myLocation):
    if (myLocation<0):
        raise MyPersonalError("I am mighty. Hear my roar!")
    print(myLocation)
```

```
locater(-1)
```

(U) When catching an exception and raising a different one, both exceptions will be raised (as of Python 3.3).

```
class MyException(Exception):
    pass

try:
    int("abc")
except ValueError:
    raise MyException("You can't convert text to an integer!")
```

(U) You can override this by adding the syntax `from None` to the end of your `raise` statement.

```
class MyException(Exception):
    pass

try:
    int("abc")
except ValueError:
    raise MyException("You can't convert text to an integer!") from None
```

(U) Testing

(U) There are two built-in modules that are pretty useful for testing your code. This also allows code to be tested each time it is imported so that a user on another machine would notice if certain methods did not do what they were intended to ahead of time.

(U) The `doctest` Module

(U) The **doctest** module allows for testing of code and value assertions in the documentation of the code itself. It also works with exceptions; you just copy and paste the appropriate `Traceback` that is expected (just the first line and the actual exception string are needed). You may incorporate `doctest` into a module or script. See the official Python documentation for details.

Doc ID: 6689693

```
"""
This is the "example" module.

The example module supplies one function, factorial().  For example,

>>> factorial(5)
120
"""

def factorial(n):
    """Return the factorial of n, an exact integer >= 0.

    >>> [factorial(n) for n in range(6)]
    [1, 1, 2, 6, 24, 120]
    >>> factorial(30)
    265252859812191058636308480000000
    >>> factorial(-1)
    Traceback (most recent call last):
        ...
    ValueError: n must be >= 0

    Factorials of floats are OK, but the float must be an exact integer:
    >>> factorial(30.1)
    Traceback (most recent call last):
        ...
    ValueError: n must be exact integer
    >>> factorial(30.0)
    265252859812191058636308480000000

    It must also not be ridiculously large:
    >>> factorial(1e100)
    Traceback (most recent call last):
        ...
    OverflowError: n too large
    """

    import math
    if not n >= 0:
        raise ValueError("n must be >= 0")
    if math.floor(n) != n:
```

```
        raise ValueError("n must be exact integer")
    if n+1 == n:  # catch a value like 1e300
        raise OverflowError("n too large")
    result = 1
    factor = 2
    while factor <= n:
        result *= factor
        factor += 1
    return result


if __name__ == "__main__":
    import doctest
    doctest.testmod()
```

(U) This lesson can be tricky to understand from the notebook. It will make the most sense if you copy and paste the above code into a file named `factorial.py`, then from the terminal run:

```
python factorial.py -v
```

Note that you don't have to include the doctest lines in your code. If you remove them, the following should work:

```
python -m doctest -v factorial.py
```

(U) The `unittest` Module

(U) The **unittest** module is much more stuctured, allowing for the developer to create a class of tests that are run and analyzed flexibly. To create a unit test for a module or script:

- `import unittest`,
- create a test class as a subclass of the `unittest.TestCase` type,
- add tests as methods of this class, making sure that the **name of each test function begins with the word 'test'**, and
- add `unittest.main()` to your main loop to run the tests.

```
import unittest
# ... other imports, script code, etc. ...
class FactorialTests(unittest.TestCase):
    def testSingleValue(self):
        self.assertEqual(factorial(5), 120)

    def testMultipleValues(self):
      ' self.assertRaises(TypeError, factorial, [1,2,3,4])

    def testBoolean(self):
        self.assertTrue(factorial(5) == 120)

def main():
    """ Main function for this script """
    unittest.main()  # Check the documentation for more verbosity levels, etc.
    # ... rest of main function ...

if __name__ == "__main__":
    main()
```

```
import unittest
dir(unittest.TestCase)
```

(U) Profiling

(U) There are many profiling modules, but we will demonstrate the **cProfile** module from the standard library. To use it interactively, first import the module, then call it with a single argument, which must be a string that could be executed if it was typed into the interpreter. Frequently, this will be a previously-defined function.

```
import cProfile
```

Doc ID: 6689693

```
def long(upper_limit=100000):
    for x in range(upper_limit):
        pass
def short():
    pass
def outer(upper_limit=100000):
    short()
    short()
    long()
```

```
cProfile.run('outer()')
```

```
cProfile.run('outer(10000000)')
```

(U) The output shows

```
ncalls: the number of calls,
tottime: the total time spent in the given function (and excluding time made in calls to sub-functions),
percall: the quotient of tottime divided by ncalls
cumtime: the total time spent in this and all subfunctions (from invocation till exit). This figure is accurate even for recu
rsive functions.
percall: the quotient of cumtime divided by primitive calls
filename:lineno(function): provides the respective data of each function
```

(U) The quick and easy way to profile a whole application is just to call the **cProfile** main function with your script as an additional argument:

```
$ python -m cProfile myscript.py
```

(U) Another useful built-in profiler is **timeit**. It's well suited for quick answers to questions like "Which is better between A and B?"

```
$ python -m timeit "'for i in range(100):' ' str(i)'
```

```
import timeit
```

```
timeit.timeit('"-".join(str(n) for n in range(100))',number=20000)
```

```
mySetup = '''
def myfunc(upper_limit=100000):
    return range(upper_limit)
'''
timeit.timeit('myfunc()',number=1000,setup=mySetup)
```

Exercise 1: Write a custom error and raise it if RangeQuery is created with dates not in the correct format.

Exercise 2: Given the list of tuples: [("2016-12-01", "2016-12-06"),("2015-12-01", "2015-12-06"),("2016-2-01", "2016-2-06"),("01/03/2014", "02/03/2014"), ("2016-06-01", "2016-10-06")] write a loop to print a rangeQuery for each of the date ranges using "~~TS//SI//REL TO USA, FVEY~~", "Primary IP address of Zendian diplomat", "10.254.18.162" as your classification, justification and selector.

Inside the loop, write a try/except block to catch your custom error for incorrectly formated dates.

Lesson 10: Iterators, Generators and Duck Typing

(b)(3)-P.L. 86-36

0

Updated 9 months ago by [redacted] in COMP 3321

3 556 273

fcs6 python

(U) Iterators, generators, sorting, and duck typing in Python.

Recommendations

UNCLASSIFIED

(U) Introduction: List Comprehensions Revisited

(U) We begin by reviewing the fundamentals of lists and list comprehension.

```
melist = [ i for i in range(1, 100, 2) ]
for i in melist:  # how does the loop work?
   print(i)
```

(U) What happens when the list construction gets more complicated?

```
noprimes = [ j for i in range(2, 19) for j in range(i*2, 500, i) ]
```

```
primes = [ x for x in range(2, 500) if x not in noprimes ]
```

```
print(sorted(primes))
```

(U) Can we do this in one shot? Yes, but...

```
# nesting madness !
primes = [ x for x in range(2, 500) if x not in [ j for i in range(2, 19) for j in range(i*2, 500, i) ] ]
```

(U) Iterators

(U) To create your own iterable objects, suitable for use in `for` loops and list comprehensions, all you need to do is implement the right special methods for the class. The `__iter__` method should return the iterable object itself (almost always `self`), and the `__next__` method defines the values of the iterator.

(U) Let's do an example, sticking with the theme previously introduced, of an iterator that returns numbers in order, except for multiples of the arguments used at construction time. We'll make sure that it terminates eventually by raising the `StopIteration` exception whenever it gets to `200`. (This is a great example of an *exception* in Python that is not uncommon: handling an event that is not unexpected, but requires termination; `for` loops and list comprehensions *expect* to get the `StopIteration` exception as a signal to stop processing.)

```
class NonFactorIterable(object):
    def __init__(self, *args):
        self.avoid_multiples = args
        self.x = 0
    def __next__(self):
        self.x += 1
        while True:
            if self.x > 200:
                raise StopIteration
            for y in self.avoid_multiples:
                if self.x % y == 0:
                    self.x += 1
                    break
            else:
                return self.x
    def __iter__(self):
        return self
```

```
silent_fizz_buzz = NonFactorIterable(3, 5)
```

```
[x for x in silent_fizz_buzz]
```

```
mostly_prime = NonFactorIterable(2, 3, 5, 7, 11, 13, 17, 19)
```

```
partial_sum = 0
```

```
for x in mostly_prime:
    partial_sum += x
```

```
partial_sum
```

```
mostly_prime = NonFactorIterable(2, 3, 5, 7, 11, 13, 17, 19)
print(sum(mostly_prime))
```

(U) It may seem strange that the `__iter__` method doesn't appear to do anything. This is because in some cases the **iterator** for an object should not be the same as the object itself. Covering such usage is beyond the scope of the course.

(U) There is another way of implementing a custom iterator: the `__getitem__` method. This allows you to use the square bracket `[]` notation for getting data out of the object. However, you still must remember to raise a `StopIteration` exception for it to work properly in for loops and list comprehensions.

Another iterator example

In the below example, we create an iterator that returns the squares of numbers. Note that in the `__next__` method, all we're doing is iterating our counter (`self.x`) and returning the square of that counter number, as long as the counter is not greater than the pre-defined limit (`self.limit`). The `while` loop in the previous example was specific to that use-case; we don't actually need to implement any looping at all in `__next__` , as that's simply the method called for each iteration through a loop on our iterator.

Here we're also implementing the `__getitem__` method, which allows us to retrieve a value from the iterator at a certain index location. This one simply calls the iterator using `self.__next__` until it arrives at the desired index location, then returns that value.

```
class Squares(object):

    def __init__(self, limit=200):
        self.limit = limit
        self.x = 0

    def __next__(self):
        self.x += 1
        if self.x > self.limit:
            raise StopIteration
        return (self.x-1)**2

    def __getitem__(self, idx):
        # initialize counter to 0
        self.x = 0
        if not isinstance(idx, int):
            raise Exception("Only integer index arguments are accepted!")
        while self.x < idx:
            self.__next__()
        return self.x**2

    def __iter__(self):
        return self
```

```
my_squares = Squares(limit=20)
```

```
[x for x in my_squares]
```

```
my_squares[5]
```

```
# since we set a limit of 20, we can't access an index location higher than that
my_squares[25]
```

(U) Benefits of Custom Iterators

1. (U) Cleaner code
2. (U) Ability to work with infinite sequences
3. (U) Ability to use built-in functions like `sum` that work with iterables
4. (U) Possibility of saving memory (e.g. `range`)

(U) Generators

(U) Generators are iterators with a much lighter syntax. Very simple generators look just like list comprehensions, except they're surrounded with parentheses `()` instead of square brackets `[]`. More complicated generators are defined like functions, with the one difference being that they use the `yield` keyword instead of the `return` keyword. A generator maintains state in between times when it is called; execution resumes starting immediately after the `yield` statement and continues until the next `yield` is encountered.

```
y = (x*x for x in range(30))
print(y)  # hmm...
```

```
def xsquared():
    for i in range(30):
        yield i*i
```

```
def xsquared_inf():
    x = 0
    while True:
        yield x*x
        x += 1
```

```
squares = [x for x in xsquared()]
print(squares)
```

(U) Another example...days of the week!

```
def day_of_week():
    i = 0
    days = ["Monday","Tuesday","Wednesday","Thursday","Friday","Saturday","Sunday"]
    while True:
        yield days[i%7]
        i += 1
day_of_week()
```

```
import random
def snowday(prob=.01):
    r = random.random()
    if r < prob:
        return "snowday!"
    else:
        return "regular day."
```

```
n = 0
for x in day_of_week():
    today = snowday()
    print(x + " is a " + today)
    n += 1
    if today == "snowday!":
        break
```

```
weekday = (day for day in day_of_week())
```

```
next(weekday)
```

(U) Pipelining

(U) One powerful use of generators is to connect them together into a *pipeline*, where each generator is used by the next. Since Python evaluates generators "lazily," i.e. as needed, this can increase the speed and potentially allow steps to run concurrently. This is especially useful if one or two steps can take a long time (e.g. a database query). Without generators, the long-running steps will become a bottleneck for execution, but generators allow other steps to proceed while waiting for the long-running steps to finish.

```
import random

# Get the fractional part of a string representation of a float
def frac_part(v):
  v = str(v)
  i, f = v.split('.')
  return f
```

```
# traditional approach
results = []
for i in range(20):
    r = random.random() *100      # generate a random number
    r_str = str(r)                # convert it to a string
    r_frac = frac_part(r_str)     # get the fractional part
    r_out = float('0.' + r_frac)  # convert it back to a float
    results.append(r_out)

results
```

```
# generator pipeline
rand_gen = ( random.random() * 100 for i in range(20) )
str_gen  = ( str(r) for r in rand_gen )
frac_gen = ( frac_part(r) for r in str_gen )
out_gen  = ( float('0.'+r) for r in frac_gen )

results = list(out_gen)
results
```

(U) Sorting

(U) In Python 2, anything iterable can be sorted, and Python will happily sort it for you, even if the data is of mixed types--by default, it uses the built-in `cmp` function, which almost always does something (except with complex numbers). However, the results may not be what you expect!

(U) In Python 3, iterable objects must have the `__lt__` (`lt` = less than) function explicitly defined in order to be sortable.

(U) The built-in function `sorted(x)` returns a new list with the data from `x` in sorted order. The `sort` method (for `list`s only) sorts a list in-place and returns `None`.

```
int_data = [10, 1, 5, 4, 2]
```

```
sorted(int_data)
```

```
int_data
```

```
int_data.sort()
```

```
int_data
```

(U) To specify how the sorting takes place, both `sorted` and `sort` take an optional argument called `key`. `key` specifies a *function* of one argument that is used to extract a comparison key from each list element (e.g. `key=str.lower`). The default value is None (compare the elements directly).

```
users = ['hAcker1', 'TheBoss', 'botman', 'turingTest']
```

```
sorted(users)
```

```
sorted(users, key=str.lower)
```

(U) The `__lt__` function takes two arguments: `self` and another object, normally of the same type.

```
class comparableCmp(complex):
    def __lt__(self, other):
        return abs(self) < abs(other)
```

```
a = 3+4j
```

```
b = 5+12j
```

```
a < b
```

```
a1 = comparableCmp(a)
```

```
b1 = comparableCmp(b)
```

```
a1 < b1
```

```
c = [b1, a1]
```

```
sorted(c)
```

(U) Here's how it works:

1. the argument given to `key` must be a function that takes a single argument;
2. internally, `sorted` creates function calls `key(item)` on each item in the list and then
3. sorts the original list by using `__lt__` on the results of the `key(item)` function.

(U) Another way to do the comparison is to use `key` :

```
def magnitude_key(a):
    return (a*a.conjugate()).real
```

```
magnitude_key(3+4j)
```

```
sorted([5+3j, 1j, -2j, 35+0j], key=magnitude_key)
```

(U) In many cases, we must sort a list of dictionaries, lists, or even objects. We could define our own key function or even several key functions for different sorting methods:

```
list_to_sort = [{'lname':'Jones', 'fname':'Sally'}, {'lname':'Jones', 'fname':'Jerry'}, {'lname':'Smith', 'fname':'John'}, {'lr
```

```
def lname_sorter(list_item):
    return list_item['lname']
```

```
def fname_sorter(list_item):
    return list_item['fname']
```

```
def lname_then_fname_sorter(list_item):
    return (list_item['lname'], list_item['fname'])
```

```
sorted(list_to_sort, key=lname_sorter)
```

```
sorted(list_to_sort, key=fname_sorter)
```

```
sorted(list_to_sort, key=lname_then_fname_sorter)
```

(U) While it's good to know how this works, this pattern common enough that there is a method in the standard library `operator` package to do it even more concisely.

```
import operator
```

```
lname_sorter = operator.itemgetter('lname') # same as previous lname_sorter
```

(U) The application of the `itemgetter` method returns a *function* that is eqivalent to the `lname_sorter` function above. Even better, when passed multiple arguments, it returns a tuple containing those items in the given order. Moreover, we don't even need to give it a name first, it's fine to do this:

```
sorted(list_to_sort, key=operator.itemgetter('lname'))
```

```
sorted(list_to_sort, key=operator.itemgetter('lname', 'fname')) # same as using lname_then_fname_sorter
```

(U) To use `operator.itemgetter` with `list` s or `tuple` s, give it integer indices as arguments. The equivalent function for objects is `operator.attrgetter` .

(U) Since we know so much about Python now, it's not hard to figure out how simple `operator.itemgetter` actually is; the following function is essentially equivalent:

```
def itemgetter_clone(*args):
    def f(item):
        return tuple(item[x] for x in args)
    return f
```

(U) Obviously, `operator.itemgetter` and `itemgetter_clone` are not actually simple--it's just that most of the complexity is hidden inside the Python internals and arises out of the fundamental data model.

(U) Duck Typing

(U) All the magic methods we've discussed are examples of the fundamental Python principle of **duck typing**: "If it walks like a duck and quacks like a duck, it must be a duck." Even though Python has `isinstance` and `type` methods, it's considered poor form to use them to validate input inside a function or method. If verification needs to take place, it should be restricted to verifying required behavior using `hasattr`. The benefit of this approach can be seen in the built-in `sum` function.

```
help(sum)
```

(U) **Any** sequence of numbers, regardless of whether it's a `list`, `tuple`, `set`, generator, or custom iterable, can be passed to `sum`.

(U) The following is a comparison of *bad* and *good* examples of how to write a `product` function:

```
def list_prod(to_multiply):
    if isinstance(to_multiply, list): # don't do this!
        accumulator = 1
        for i in to_multiply:
            accumulator *= i
        return accumulator
    else:
        raise TypeError("Argument to_multiply must be a list")
```

```
def generic_prod(to_multiply):
    if hasattr(to_multiply, '__iter__') or hasattr(to_multiply, '__getitem__'):
        accumulator = 1
        for i in to_multiply:
            accumulator *= i
        return accumulator
    else:
        raise TypeError("Argument to_multiply must be a sequence")
```

```
list_prod([1,2,3])
```

```
list_prod((1,2,3))
```

```
generic_prod((1,2,3))
```

(U) Having given that example, testing for iterability is one of a few special cases where `isinstance` might be the right function to use, but not in the obvious way. The `collections` package provides **abstract base classes** which have the express purpose of helping to determine when an object implements a common interface.

(U) Finally, effective use of duck typing goes hand in hand with robust error handling, based on the principle that "it's easier to ask for forgiveness than permission."

Exercises

1. Add a method to your 'RangedQuery' class to allow instances of the class to be sorted by 'start_date'.
2. Write an iterator class 'ReverseIter' that takes a list and iterates it from the reverse direction.
3. Write a generator which will iterate over every day in a year. For example, the first output would be 'Monday, January 1'.
4. Modify the generator from exercise 2 so the user can specify the year and initial day of the week.

Pipelining with Generators

0

Created over 3 years ago by [redacted]

Python3 thumbnail 135 20

fcs6 data generators looping transformation pipeline pipelining processing laundry

(U) Defining processing pipelines with generators in Python. It's simply awesome.

Recommendations

Pipelining with Generators

Imagine you're doing your laundry. Think about the stages involved. Roughly speaking, the stages are sorting, washing, drying, and folding. The beauty though is that even though these stages are sequential, they can be performed in parallel. This is called **pipelining**.

Python generators make pipelining easy and can even clarify your code quite a bit. By breaking your processing into distinct stages, the Python interpreter can make better use of your computer's resources, and even break the stages out into separate threads behind the scenes. Memory is also conserved because values are automatically generated as needed, and discarded as soon as possible.

A prime example of this is processing results from a database query. Often, before we can use the results of a database query, we need to clean them up by running them through a series of changes or transformations. Pipelined generators are perfect for this.

```
from pprint import pprint
import random
```

A Silly Example

Here we're going to take 200 randomly generated numbers and extract their fractional parts (the part after the decimal point). There are probably more efficient ways to do this, but we're doing to do it by splitting out the string into two parts. Here we have a function that simply returns the integer part and the fractional part of an input float as two strings in a tuple.

```
def split_float(v):
    '''
    Takes a float or string of a float
    and returns a tuple containing the
    integer part and the fractional part
    of the number, as strings, respectively.
    '''
    v = str(v)
    i, f = v.split('.')
    return (i, '0.'+f)
```

The Pipeline

Here we have a pipeline of four generators, each feeding the one below it. We pprint out the final resulting list after all the stages have complete. See the comments after each line for further explanation.

```
rand_gen = ( random.random() * 100 for i in range(200) )  # generate 200 random floats between 0 and 100, one at a time
results = ( split_float(r) for r in rand_gen )  # call our split_float() function which will generate the correspoding tuples
results = ( r[1] for r in results )  # we only care about the fractional part, so only keep that part of the tuple
results = ( float(r) for r in results )  # convert our fractional value from a string back into a float

pprint(list(results))  # print the final results
```

Why not a for-loop?

We could have put all the steps of our pipeline into a single for-loop, but we get a couple advantages by breaking the stages out into separate generators:

- There's some clarity gained by having distinct stages specified as a pipeline. People reading the code can clearly see the transforms.
- In a for-loop, Python simply computes the values sequentially; there's no chance for automatic optimization or multi-threading. By breaking the stages out, each stage can execute in parallel, just like your washer and dryer.

Another (Pseudo-)Example

Here's a pseudo-example querying a database that returns JSON that we need to convert to lists.

```
import json
results = ( json.loads(result) for result in db_cursor.execute(my_query) )
results = ( r['results'] for r in results )
results = ( [ r['name'], r['type'], r['count'], r['source'] ] for r in results )
```

Filters

We can even filter our data in our generator pipeline.

```
results = ( r for r in results if r[2] > 0 )   # remove results with a count of zero
foo(results)  # do something else with your results
```

Lesson 11: String Formatting

(b)(3)-P.L. 86-36

0

Updated 9 months ago by [redacted] in COMP 3321

3 1 547 251

python fcs6

(U) Lesson 11: String Formatting

Recommendations

UNCLASSIFIED

(U) Intro to String Formatting

(U) String formatting is a very powerful way to display information to your users and yourself. We have used it through many of our examples, such as this:

```
'This is a formatted String {}'.format("--->hi I'm a formatted String argument<---")
```

(U) This is probably the easiest example to demonstrate. The empty curly brackets `{}` take the argument passed into `format`.

(U) Here's a more complicated example:

```
'{2} {1} and {0}'.format('Henry', 'Bill', 'Bob')
```

(U) Arguments can be positional, as illustrated above, or named like the example below. Order does matter, but names can help.

```
'{who} is really {what}!'.format(who='Tony', what='awesome')
```

(U) You can also format lists:

```
cities = ['Dallas', 'Baltimore', 'DC', 'Austin', 'New York']

'{0[4]} is a really big city.'.format(cities)
```

(U) And dictionaries:

```
lower_to_upper = {'a':'A', 'b':'B', 'c':'C'}

"This is a big letter {0[a]}".format(lower_to_upper)  # notice no quotes around a
```

```
"This is a big letter {lookup[a]}".format(lookup=lower_to_upper)  # can be named
```

```
for little, big in lower_to_upper.items():
    print('[-->{0:10} -- {1:10}<--]'.format(little, big))
```

(U) If you actually want to include curly brackets in your printed statement, use double brackets like this: {{ }} .

```
"{{0}} {0}".format('Where do I get printed?')
```

(U) You can also store the format string in a variable ahead of time and use it later:

```
the_way_i_want_it = '{0:>6} = {0:>#16b} = {0:#06x}'

for i in 1, 25, 458, 7890:
    print(the_way_i_want_it.format(i))
```

(U) Format Field Names

(U) Here are some examples of field names you can use in curly brackets within a format string.

{<field name>}

- (U) 1 : the *second* positional argument
- (U) name : keyword argument
- (U) 0.var : attribute named var of the first positional argument
- (U) 3[0] : element 0 of the fourth positional argument
- (U) me_data[key] : element associated with the specific key string 'key' of me_data

(U) Format Specification

(U) When using a format specification, it follows the field name within the curly brackets, and its elements must be in a certain order. This is only for reference; for a full description, see the Python documentation on string formatting.

```
{<field name>:<format spec>}
```

1. (U) Padding and Alignment

- `>` : align right
- `<` : align left
- `=` : only for numeric types
- `^` : center

1. (U) Sign

- `-` : prefix negative numbers with a minus sign
- `+` : like `-` but also prefix positive numbers with a `+`
- `' '` : like `-` but also prefix positive numbers with a space

1. (U) Base Indicator (precede with a hash `#` like above)

- `0b` : binary
- `0o` : octal
- `0x` : hexadecimal

1. (U) Digit Separator

- `,` : use a comma to separate thousands

1. (U) Field Width

- leading `0` : pad with zeroes at the front

1. (U) Field Type (letter telling which type of value should be formatted)

- `s` : string (the default)
- `b` : binary
- `d` : decimal: base 10
- `o` : octal
- `x` : hex uses lower case letters
- `X` : hex uses upper case
- `n` : like `d`, use locale settings to determine decimal point and thousands seperator
- no code integer : like `d`
- `e` : exponential with small e
- `E` : exponential with big E
- `f` : fixed point, `nan` for not a number and `inf` for infinity
- `F` : same as `f` but uppercase `NAN` and `INF`

- `g` : general format
- `G` : like `G` but uppercase
- `n` : locale settings like `g`
- `%` : times 100, displays as `f` with a `%`
- no code decimal : like `g` , precision of twelve and always one spot after decimal point

1. (U) Variable Width

(U) New in Python 3.6: f-strings

```
# Add 'f' before the string to create an f-string
# Expression added directly inside the `{}` brackets rather than after the format statement
x = 34
y = 2
f"34 * 2 = {x*y}"
```

```
my_name = 'Bob'
f"My name is {my_name}"
```

(U) Examples

```
'{0:{1}.{2}f}'.format(9876.5432, 18, 3)
```

```
'{0:010.4f}'.format(-123.456)
```

```
'{0:+010.4f}'.format(-123.456)
```

```
for i in range(1, 6):
    print('{0:10.{1}f}'.format(123.456, i))
```

```
v = {'value':876.543, 'width':15, 'precision':5}

"{0[value]:{0[width]}.{0[precision]}}".format(v)
```

```
data = [('Steve', 59, 202), ('Samantha', 49, 156), ('Dave', 61, 135)]

for name, age, weight in data:
    print('{0:<12s} {1:4d} {2:4d}'.format(name, age, weight))
```

```
# same as above but with f-strings

data = [('Steve', 59, 202), ('Samantha', 49, 156), ('Dave', 61, 135)]

for name, age, weight in data:
    print(f'{name:<12s} {age:4d} {weight:4d}')
```

COMP3321 Day01 Homework - GroceryList

0

(b)(3)-P.L. 86-36

Updated almost 3 years ago by [redacted] in COMP 3321

3 233 36

python exercises

(U) Homework for Day01 of COMP3321. Task is to sort items into bins.

Recommendations

(U) COMP3321 Day01 Homework GroceryList

Doc ID: 6689693

```
myGroceryList = ["apples", "bananas", "milk", "eggs", "bread",
                 "hamburgers", "hotdogs", "ketchup", "grapes",
                 "tilapia", "sweet potatoes", "cereal",
                 "paper plates", "napkins", "cookies",
                 "ice cream", "cherries", "shampoo"]

## Items by category
vegetables = ["sweet potatoes", "carrots", "broccoli", "spinach",
              "onions", "mushrooms", "peppers"]
fruit = ["bananas", "apples", "grapes", "plumbs", "cherries", "pineapple"]
cold_items = ["eggs", "milk", "orange juice", "cheese", "ice cream"]
proteins = ["turkey", "tilapia", "hamburgers", "hotdogs", "pork chops", "ham", "meatballs"]
boxed_items = ["pasta", "cereal", "oatmeal", "cookies", "ketchup", "bread"]
paper_products = ["toilet paper", "paper plates", "napkins", "paper towels"]
toiletry_items = ["toothbrush", "toothpaste", "deodorant", "shampoo", "soap"]

## My items by category
my_vegetables = []
my_fruit = []
my_cold_items = []
my_proteins = []
my_boxed_items = []
my_paper_products = []
my_toiletry_items = []
```

(U) Fill in your code below. Sort the items in myGroceryList by type into appropriate my_category lists using looping and decision making

```
print("My vegetable list: ", my_vegetables)
print("My fruit list: ", my_fruit)
print("My cold item list: ", my_cold_items)
print("My protein list: ", my_proteins)
print("My boxed item list: ", my_boxed_items)
print("My paper product list: ", my_paper_products)
print("My toiletry item list: ", my_toiletry_items)
```