Theory: Reading files

🕓 21 minutes 7 / 7 problems solved

Start practicing

4054 users solved this topic. Latest completion was about 1 hour ago.

If you have a file, one thing you might want to do with it is seeing its contents. In this topic, you'll learn different ways how to read files in Python.

§1. Read file

To read the file first we need to open it in the reading mode. Now that the file is opened, how does the actual reading part go? Let's first choose a file that we want to read. So, imagine we have a file called '**animals.txt**' that looks like this:

```
Dog
Cat
Rabbit
Sea turtle
Penguin
```

To read the file you can:

- use the `read(size)` method;
- use the `readline(size)` method;
- use the `readlines()` method;
- iterate over the lines with a `for` loop.

The first three ways are special file object methods while the last one is a general Python loop. Let's go over them one by one.

§2. read()

`read(size)` reads the `size` bytes of a file. If the parameter isn't specified, the whole file is read into a single variable. So, this is what we'll get if we apply it to our file:

```
file = open('animals.txt', 'r')
print(file.read())
# The output:
# Dog
# Cat
# Rabbit
# Sea turtle
# Penguin

file.close()
```

§3. readline()

`readline(size)` is similar to `read(size)` but it reads `size` bytes from a single line, not the whole file. Lines in files are separated by **newline** escape sequences: `'\n'`, `'\r'` or `'\r\n'`. We'll settle on `'\n'` in this topic. Yet, keep in mind that this escape sequence depends on your operating system.

Let's proceed with our example. The file '**animals.txt**' contains 5 lines. Here is what we'll get if we try to read 3 bytes from each line:

Current topic:

✓ Reading files Stage 2 ···

Topic depends on:

✓ For loop 12★ Stage 1 ···

✓ Files in Python Stage 2 ···

Topic is required for:

✓ Context manager Stage 2 ···

Built-in exceptions ···

Table of contents:

```
file = open('animals.txt', 'r')
print(file.readline(3))
print(file.readline(3))
print(file.readline(3))
print(file.readline(3))
print(file.readline(3))
print(file.readline(3))

# The output:
# Dog
#
#
# Cat
#
#
# Rab
# bit


file.close()
```

As you can see, the output does not contain the first three characters from all 5 lines. This is because when we specify the `size` parameter, we get `size` bytes from the line until it ends and only then go to another line. The newline character is considered as a part of the line here. So, in our example, we need two passes of `readline(3)` to read the first line which has three characters and a newline character (4 bytes in total). This is where all those empty lines in the output come from.

§4. readlines()

`readlines()` allows us to read the whole file as a list of lines. Here's what it looks like:

```
file = open('animals.txt', 'r')
print(file.readlines())
# The output:
# ['Dog\n', 'Cat\n', 'Rabbit\n', 'Sea turtle\n', 'Penguin']

file.close()
```

Here you can actually see that the newline character is a part of the line! This method works well when the files are small, but for large files, this may be inefficient.

§5. for loop

The most efficient way to read the contents of a file is to iterate over its lines with `for` loop.

```python
file = open('animals.txt', 'r')
for line in file:
    print(line)

# The output:
# Dog
#
# Cat
#
# Rabbit
#
# Sea turtle
#
# Penguin

file.close()
```

This is the best way to read large files because we can just work with one line at a time or work with specific lines and discard the rest.

§6. Summary

So, in this topic, we've seen how to read the file's contents. This skill is quite useful because sometimes it's easier or more convenient to read a file using Python than to open it directly. Again, these are the available ways to do it:

- `read()` method reads the file as a whole;
- `readline()` reads the file one line at a time;
- `readlines()` reads the file as a list of lines.
- the best way, however, is to use `for` loop to iterate over the lines of the file.

Report a typo

Thanks for your feedback!

Write here how we could improve this theory

Start practicing

Comments (16) Hints (3) Useful links (0) Show discussion