Memory and Disk Usage: Takeaways

by Dataquest Labs, Inc. - All rights reserved © 2020

Syntax

- Creating a fixed bit-length integer:

```
import numpy as np
x = np.int8(100)
```

- Obtaining the binary representation of a fixed bit-length integer:

```
np.binary_repr(2147483647, width=32)
```

- Obtaining the size in bytes of a Python object in memory:

```
import sys
sys.getsizeof([0, 4, 2])
```

- Obtaining the numbers of bits necessary to represent a number:

```
int.bit_length(42)
```

- Obtaining the size, in bytes, of a file:

```
import os
os.path.getsize('filename')
```

Concepts

- Two's complement representation makes it possible to represent both positive and negative numbers using a fixed number of bits.
- The values of the digits in the two's complement representation is the same except the leftmost is negated. For example, with 5 bits, the weights are (from left to right): -16, 8, 4, 2, 1.
- Overflow occurs when the result of a computation would yield a number that is outside of the range that those numbers can represent.
- With n bits, two's complement representation can represent integers from -2^{n-1} to $2^{n-1} - 1$.
- In Python, strings are stored using one of the following encodings: Latin-1 (1 byte), UCS-2 (2 bytes) or UCS-4 (4 bytes). The one that is used depends on the characters in the string. All characters in a string are encoded using the same encoding.
- Different encodings occupy different space on disk.

Resources

- Integer data types
- numpy types
- Two's complement representation