String vs. Bytes

Test in Python 3 is always unicode and is represented by the **str** type, and binary data is represented by the **bytes** type. They **cannot** be mixed.

Strings can be **encoded** to bytes, and bytes can be **decoded** back to strings.

```
In [1]: s = 'Hello world!'
        print(s)
        print("Length is", len(s))

        Hello world!
        Length is 12
```

```
In [2]: us = 'Hello 世界!'
        print(us)
        print("Length is", len(us))

        Hello 世界!
        Length is 9
```

```
In [3]: bs = s.encode('utf-8')
        print(bs)
        print("length is", len(bs))

        b'Hello world!'
        length is 12
```

```
In [4]: bus = us.encode('utf-8')
        print(bus)
        print("length is", len(bus))

        b'Hello \xe4\xb8\x96\xe7\x95\x8c!'
        length is 13
```

```
In [5]: print(bs.decode('utf-8'))
        print(bus.decode('utf-8'))

        Hello world!
        Hello 世界!
```